

VYSOKÁ ŠKOLA BÁŇSKÁ – TECHNICKÁ UNIVERZITA OSTRAVA
EKONOMICKÁ FAKULTA

KATEDRA APLIKOVANÉ INFORMATIKY

Vytvoření dashboardu na základě analýzy dat z informačního systému
Dashboard Creation Based on a Data Analysis from the Information System

Student: Bc. Adam Rybář
Vedoucí diplomové práce: Ing. Pavel Vlček, Ph.D.

Ostrava 2015

Zadání diplomové práce

Student: **Bc. Adam Rybář**

Studijní program: N6209 Systémové inženýrství a informatika

Studijní obor: 1802T001 Aplikovaná informatika

Téma: Vytvoření dashboardu na základě analýzy dat z informačního systému
Dashboard Creation Based on a Data Analysis from the Information System

Zásady pro vypracování:

1. Úvod
2. Teoretická východiska a použité technologie
3. Analýza současného stavu dat ve vybrané firmě
4. Možnosti a návrh řešení vybraným SW nástrojem
5. Závěr

Seznam použité literatury

Seznam zkratk

Prohlášení o využití výsledků diplomové práce

Seznam příloh

Přílohy

Seznam doporučené odborné literatury:

LARSON, Brian. *Delivering Business Intelligence with Microsoft SQL Server 2012*. 3. vyd. New York: McGraw-Hill, 2012. ISBN 978-0071759380.

STEPHENS, Ryan, Ronald PLEW a Arie JONES. *Naučte se SQL za 28 dní*. Brno: Computer Press, 2010. ISBN 978-80-251-2700-1.

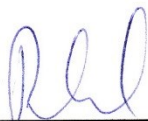
WINSTON, Wayne L. *Microsoft Excel 2013: Data Analysis and Business Modeling*. Seattle: Microsoft Press, 2014. ISBN 978-0-7356-6913-0.

Formální náležitosti a rozsah diplomové práce stanoví pokyny pro vypracování zveřejněné na webových stránkách fakulty.

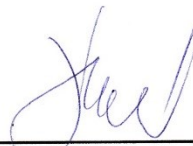
Vedoucí diplomové práce: **Ing. Pavel Vlček, Ph.D.**

Datum zadání: 21.11.2014

Datum odevzdání: 25.04.2015



Ing. Petr Rozehnal, Ph.D.
vedoucí katedry



prof. Dr. Ing. Dana Dluhošová
děkanka fakulty

Prohlášení

Prohlašuji, že jsem celou diplomovou práci, včetně všech příloh, vypracoval samostatně a veškeré zdroje jsem uvedl v seznamu literatury.

V Ostravě dne 25.04.2015

.....
Bc. Adam Rybář

Poděkování

Rád bych poděkoval svému vedoucímu, Ing. Pavlu Vlčkovi, Ph.D., za cenné rady a odbornou pomoc, díky které jsem byl schopen vytvořit tuto diplomovou práci.

Obsah

1	Úvod.....	5
2	Teoretická východiska a použité technologie	6
2.1	Business Intelligence	6
2.1.1	Business Intelligence obecně	6
2.1.2	Efektivní rozhodování	7
2.2	Základní principy řešení Business Intelligence	8
2.2.1	Výběr a organizace dat	9
2.2.2	Dimenze a granularita dat	9
2.2.3	Multidimenzionalita uložení a práce s daty.....	10
2.3	Vrstvy, typy aplikací a hlavní komponenty BI.....	11
2.4	Hlavní komponenty řešení BI.....	14
2.4.1	Produkční (zdrojové) databáze.....	14
2.4.2	Extraction, Transformation and Loading – ETL.....	15
2.4.3	Architektura datových úložišť	15
2.4.4	OLAP databáze	20
2.4.5	Reporting.....	20
2.4.6	Analytické aplikace	20
2.4.7	Dolování dat	20
2.4.8	Nástroje pro řízení kvality dat a správu metadat.....	21
2.4.9	Enterprise Application Integration – EAI	21
2.5	Úvod do analytických služeb.....	22
2.5.1	Relační databáze.....	22
2.5.2	Multidimenzionální databáze	25
2.5.3	Porovnání relačního a multidimenzionálního modelu	26
2.6	Problematika analytických databází	26
2.6.1	Fakta	27
2.6.2	Dimenze	27
2.6.3	Schémata	28
2.6.4	Úložiště multidimenzionálních dat.....	29
2.7	Reportovací služby	30
2.7.1	Úvod do reportingu	30
2.7.2	Kategorizace reportů	31

2.7.3	Životní cyklus reportu	32
2.8	Dashboard	33
2.9	Segment trhu Business Intelligence	36
3	Analýza současného stavu dat ve vybrané firmě	38
3.1	Představení společnosti	38
3.2	Současný stav ve firmě	38
4	Možnosti a návrh řešení vybraným SW nástrojem	40
4.1	Příprava zdroje dat pro softwarové nástroje	40
4.2	SQL Server Reporting Services (SSRS)	41
4.3	Power Pivot	43
4.4	Power View	45
4.5	Zvolení SW nástroje	46
4.6	Návrh a testování dashboardu	47
4.6.1	Datový model	47
4.6.2	Testovací verze dashboardu	48
4.6.3	Testování	49
4.7	Finální verze dashboardu	49
4.8	Shrnutí	54
5	Závěr	55
	Seznam použité literatury	56
	Seznam zkratk	57
	Prohlášení o využití výsledků diplomové práce	
	Seznam příloh	

1 Úvod

Informace hýbou dnešním světem, a kdo je dokáže dobře ovládat a využít má obrovskou výhodu. Na základě získaných informací děláme určitá rozhodnutí, a ta pak ovlivňují dění kolem nás. Jak se říká, klíčem k úspěšnému životu je - dělat správná rozhodnutí.

Vrcholoví manažeři musí dělat desítky rozhodnutí týdně. Jelikož člověk nemá neomezenou paměť a nedokáže pojmout veškeré potřebné informace a nakládat s nimi efektivně, začaly vznikat informační systémy (IS) na bázi informačních technologií, které jsou schopné pojmout veškeré procesy ve firmě od účetnictví, přes výrobu až po logistiku. Mimo tyto všeobslhlé informační systémy existují nástroje zaměřující se pouze na práci s daty a jejich analýzu. A právě Business Intelligence tyto analytické nástroje a aplikace využívá pro strategická rozhodování ve firmách. Podle Tvrdíkové (2008) pojem Business Intelligence (BI) znamená sběr, integraci, analýzu, interpretaci a prezentaci informací, na základě kterých můžeme řídit organizaci podle znalostí.

V době, kdy je třeba zpracovávat data co nejpřesněji a nejrychleji jsou BI nástroje obrovskou výhodou pro management, který dokáže mimo jiné pohotově reagovat na situaci odehrávající se jak uvnitř firmy tak i mimo ní.

Cílem této diplomové práce je vytvoření dashboardu pro firmu AVE Soft s.r.o., která z navrženého dashboardu bude schopna vyčíst různé ukazatele informačního systému Evolio a podle toho například ovlivnit nasazování nových verzí systému jednotlivým zákazníkům.

Pro vytvoření dashboardu budou vyzkoušeny tři nástroje: Power Pivot, Power View a SSRS. Pomocí každého nástroje bude vytvořen náhled a po srovnání bude vybrán nejvhodnější softwarový nástroj, ve kterém bude vytvořen kompletní dashboard. Všechny tři zvolené nástroje jsou pouze od firmy Microsoft, protože společnost AVE Soft s.r.o. má s firmou Microsoft bohaté zkušenosti.

Diplomová práce je rozdělena na tři části. Na teoretickou, analytickou a praktickou část. V teoretické části jsou popsána teoretická východiska a použité technologie, které jsou důležité pro část praktickou. V analytické části je věnován prostor pro popis společnosti, pro kterou je dashboard vyvíjen a pro analýzu původního řešení prezentace dat. V praktické části je řešen výběr vhodného nástroje pro vytvoření dashboardu a poté jeho vývoj v daném nástroji.

2 Teoretická východiska a použité technologie

Celá tato část se zabývá teoretickými východisky sloužícími k lepšímu pochopení praktické části. Mimo jiné bude popsáno co je to Business Intelligence a jaké jsou principy a komponenty řešení BI. Dále bude nahlédnuto do oblasti analytických a reportovacích služeb a nakonec bude popsáno co je to dashboard a z čeho se může skládat.

2.1 Business Intelligence

V této kapitole bude zmíněno co je to Business Intelligence, včetně uvedení několika definic a trendu vývoje BI v České republice. Malou vsuvkou bude podkapitola o efektivním rozhodování.

2.1.1 Business Intelligence obecně

Novotný (2005, str. 17) uvádí že: „v prostředí stále tvrdší konkurence musí podnikoví analytici a manažeři rozhodovat pod časovým tlakem a současně s vysokou zodpovědností. Proto je nutné mít pro tato rozhodnutí dostatek relevantních a objektivních informací, které jsou dostupné rychle, s minimální technickou náročností na manipulaci, a přitom s možností rychle formulovat nové požadavky na další informace odpovídající aktuální obchodní nebo výrobní situaci“.

Business Intelligence (BI) je v současnosti nejperspektivnější oblastí podnikové informatiky. Je to dáno mimo jiné díky jejím možnostem efektivní podpory řídicích aktivit podnikových manažerů a specialistů. Tím aplikace BI významnou měrou přispívají k celkovému zvyšování kvality podnikové informatiky, podnikového řízení a současně se stávají i důležitým faktorem zvyšování konkurenceschopnosti podniku.

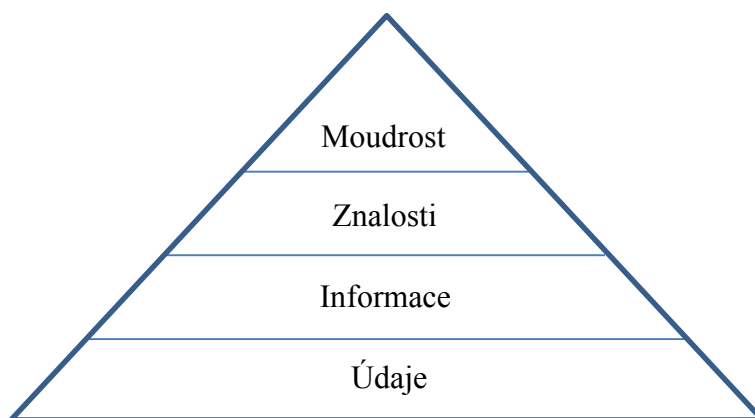
BI se v České republice rozvíjí od počátku devadesátých let a v dnešní době tento typ aplikací využívá a rozvíjí přibližně 50 % podniků a toto číslo se v posledních letech každým rokem zvyšuje.

Řešení BI v sobě zahrnuje obrovské množství nejrozličnějších organizačních, analytických, implementačních a provozních úloh využívající rozsáhlé spektrum softwarových nástrojů počínaje databázových a různých speciálně zaměřených končíc, mezi které se například řadí ETL nástroje mající na starost transformaci, kontrolu a čištění dat (Pour, a další, 2012).

Business Intelligence se podle Lacka (2009, str. 14) dá definovat také jako: „množina konceptů a metodik, které zlepšují rozhodovací proces za použití metrik, nebo systémů založených na metrikách. Účelem procesu je konvertovat velké objemy dat na poznatky,

kteře jsou potřebné pro koncové uživatele. Tyto poznatky potom lze efektivně použít například v procesu rozhodování a mohou tvořit velmi významnou konkurenční výhodu“.

Přeměnu dat na informace, informace na znalosti a budování „moudrosti“ na základě znalostí, lze zobrazit na hierarchické pyramidě informačních úrovní. Základem všeho jsou data, která obsahují jednoduchá fakta, ve kterých jsou ukryté určité informace. Ty vyplynou na povrch až tehdy, když k datům přidáme souvislosti. Když přidáme i tvořivou inteligenci, získáme znalosti a jejich zobecněním pak získáme moudrost, která bude znamenat schopnost přesného zhodnocení znalostí a jejich následné uplatnění v praxi.



Obrázek 2-1 Hierarchie informačních úrovní (Lacko, 2009, str. 15)

Často je potřeba sledovat trend nějaké veličiny, například při obchodování s cennými papíry, kdy je potřeba najít mezi údaji určité závislosti. Proto moderní databázové systémy obsahují rozsáhlou podporu pro budování datových skladů, OLAP analýzy, data mining a další (Lacko, 2009).

2.1.2 Efektivní rozhodování

Každý den je učiněno mnoho rozhodnutí, která definují směřování a efektivitu dané firmy. Může se týkat produkce, marketingu nebo třeba personálu. Rozhodování ovlivňuje náklady, prodeje, marže a stejně jako v osobním životě je klíčem úspěšné firmy dělat správná rozhodnutí.

Ta nejdůležitější, strategická, rozhodnutí, skládající se z dílčích, má na starosti top management firmy. Podřízení mají na starosti dílčí, a pokud nejsou dobře implementována, může být celé strategické rozhodnutí zcela neefektivní.

Efektivní rozhodování firmy musí obsahovat tyto tři prvky:

- cíle;
- měřitelnost, zda se plní;
- včasná zpětná vazba na základě výsledků měřitelnosti.



Obrázek 2-2 Tři klíče k efektivnímu rozhodování podle Larsona (2002, str. 6)

Každý cíl musí být vždy specifický a měl by poskytovat prostředek k jeho měření, abychom ověřili, zda se jde správným směrem. Výsledky měření musí být včas poskytnuty člověku, který je zodpovědný za určité rozhodnutí, aby mohl včas reagovat na případné odklonění se od požadovaného cíle. Otázkou tedy je, jak dané informace týkající se plnění cíle získávat a distribuovat? Odpovědí je business intelligence. Z této kapitoly nám vyplývá jedna z mnoha definic BI a to, že business intelligence znamená poskytování přesných a užitečných informací odpovědnému člověku za rozhodnutí v co nejkratším čase, tak aby dané rozhodnutí bylo efektivní (Larson, 2002).

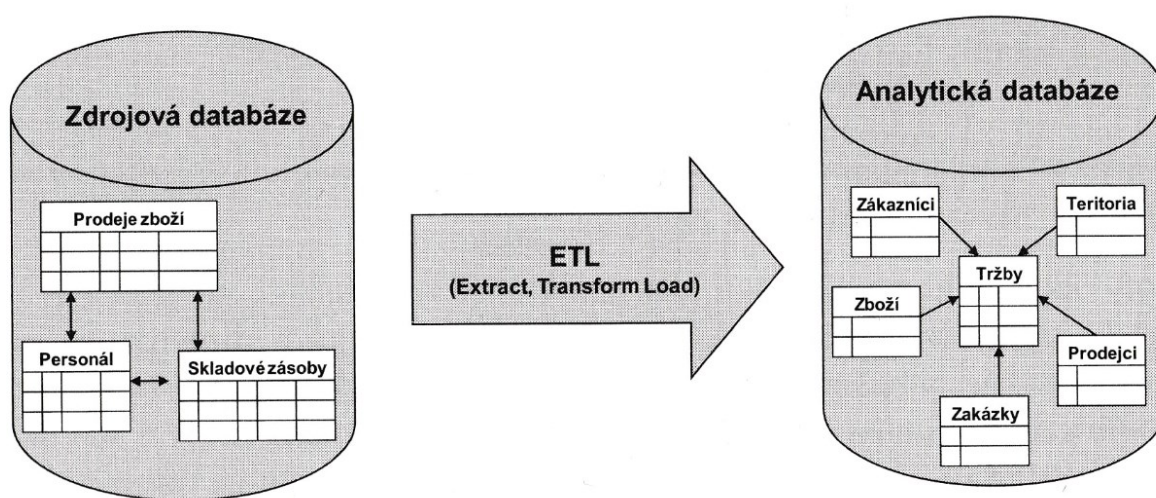
2.2 Základní principy řešení Business Intelligence

V praxi existuje nepřeberné množství variant řešení BI, které kombinují různé BI technologie a produkty. Hlavní principy Business Intelligence Pour (2012, str. 16):

- BI je určená pro analytické a plánovací aplikace a tomu odpovídá i organizace dat v jejích databázích;
- BI uchovává data na potřebných úrovních detailu, tzv. granularity, tedy detailní i agregovaná;
- BI pracuje primárně s daty podnikových ukazatelů a ty vyhodnocuje podle nejrůznějších dimenzí a jejich kombinací, je tedy založena na multidimenzionalitě uložení a zpracování dat;
- BI využívá časové dimenze, která umožňuje ukládání dat do databází postupně v časových intervalech a časových snímcích;
- BI je charakteristická vyššími nároky na kvalitu dat.

2.2.1 Výběr a organizace dat

Kromě plánovacích aplikací, BI aplikace nevytvářejí nebo nepořizují nová data, ale využívají data vytvořená transakčními aplikacemi (CRM, ERP atd.), které se z pohledu BI nazývají jako zdrojové či produkční. Podstatnou vlastností těchto databází je organizace jejich dat podporující přístupy k detailním datům, ukládání a aktualizace dat a jsou pro tyto operace optimalizované. Naopak analytické aplikace BI jsou optimalizované na efektivní poskytování analytických informací, to znamená, že musí obsahovat hodnoty ukazatelů ve vazbě na analytická hlediska, tedy dimenze. Z toho plyne, že mezi zdrojovými databázemi a analytickými databázemi musí probíhat určitá transformace dat, neboli tzv. ETL proces, kterému je vyhrazena kapitola 2.4.2 (Pour, a další, 2012).



Obrázek 2-3 Transformace dat (Pour, a další, 2012, str. 17)

2.2.2 Dimenze a granularita dat

Pour, Maryška a Novotný (2012, str. 17) uvádí, že: „dimenzí se z pohledu jejího užití rozumí analytické hledisko pro hodnocení sledovaných uživatelů, z informačního pohledu jako struktura dat, případně jako databázová tabulka obsahující záznamy o jednotlivých prvcích dimenze, například o položkách zboží apod. Prvky dimenzí, například již zmíněné zboží, jsou většinou uspořádány v hierarchické struktuře“. Například rozdělení na kategorie zboží (DVD přehrávače atd.), skupiny zboží (přenosné, stolní atd.) a na jednotlivé položky (DVD přehrávač Panasonic atd.). Počet hierarchických úrovní ve struktuře (skupin, podskupin atd.) je určen podle potřeby.

Hodnoty ukazatelů se ukládají do analytických databází na nejvyšší úrovni detailu, neboli nejvyšší granularitě, který odpovídá prvkům dimenzí na nejnižší úrovni hierarchie. Například tržby za DVD přehrávač Panasonic. Současně se do těchto databází, zejména OLAP

databází, ukládají i agregované a další vypočtené hodnoty ukazatelů, například za kategorii zboží atd.

Tato metoda slouží pro urychlení odezvy systému na analytické požadavky. Hierarchie uložení agregovaných dat umožňuje uživateli rychle a pružně se pohybovat po požadovaných úrovních agregace, bez nutnosti znovu požadované agregace počítat. Princip zpřístupnění dat na vyšší úroveň detailu se označuje jako dril-down, v opačném směru pak dril-up (Pour, a další, 2012).

2.2.3 Multidimenzionalita uložení a práce s daty

Důvodem multidimenzionality uložení a práce s daty je požadavek pohledů uživatele na sledované ukazatele z více hledisek (dimenzí) a jejich kombinací. To s sebou současně přináší i požadavek na specifickou organizaci dat v databázi. Existují dva způsoby realizace multidimenzionality v datech (Pour, 2012):

- multidimenzionalita dat v prostředí relační databáze;
- multidimenzionalita dat v prostředí OLAP technologie.

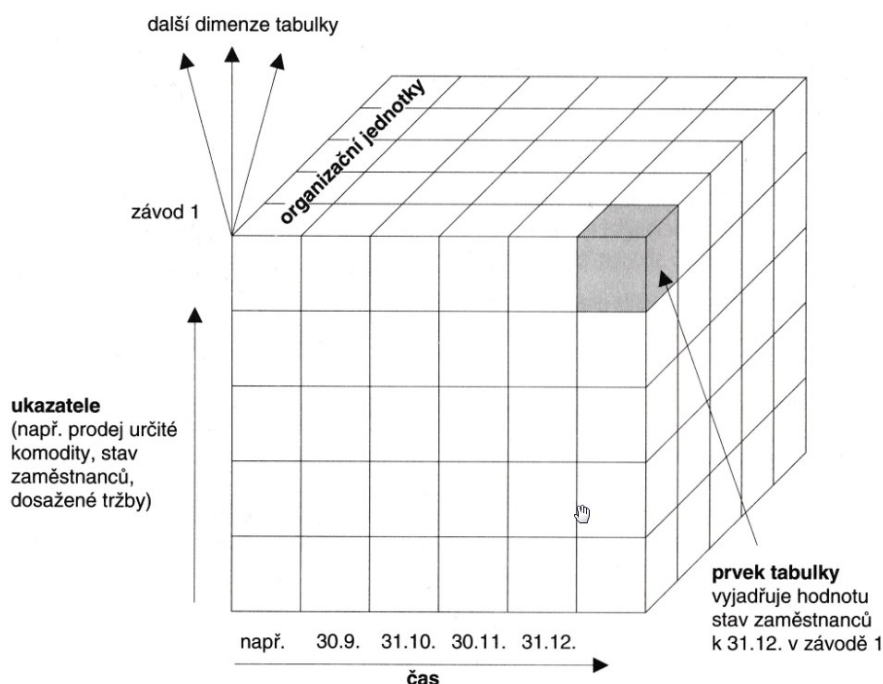
Multidimenzionalita dat v prostředí relační databáze

Relační dimenzionální model má dvě základní podoby, a to hvězdicové schéma (star scheme) a schéma sněhové vločky (snowflake scheme). Pour, Maryška a Novotný (2012, str. 17) zmiňuje, že: „v centru těchto schémat je tabulka faktů, tedy tabulka sledovaných ekonomických a dalších ukazatelů identifikovaných klíčem složených z klíčů dimenzionálních tabulek, v nichž jsou uloženy prvky jednotlivých dimenzí. Dimenzionální tabulky více méně slouží pro uložení popisných informací o hodnotách uložených v tabulce faktů. Typicky si ji lze představit jako číselník.“

Multidimenzionalita dat v prostředí OLAP technologie

Multidimenzionální databáze jsou optimalizované pro uložení a interaktivní využívání multidimenzionálních dat. Výhodou multidimenzionality, respektive nasazení OLAP technologií, je rychlost zpracování a efektivní analýzy multidimenzionálních dat, například drilling, slice and dice a podobně. Spolu se zavedením pojmu Business Intelligence a rozvojem nástrojů a technologií pro širokou podporu analytických činností v podniku se výraz OLAP více zpřesnil. Pojem OLAP se definuje jako informační technologie, založená především na koncepci multidimenzionálních databází. Hlavním principem je několikadimenzionální

tabulka umožňující rychle a pružně měnit jednotlivé dimenze a nabízet tak různé pohledy uživatele na modelovanou ekonomickou realitu.



Obrázek 2-4 Princip multidimenzionální databáze (Novotný, 2005, str. 22)

Pour, Maryška a Novotný (2012, str. 22) rovněž uvádí, že: „Obsah dimenzí je tvořen prvky dimenzí a jejich promítnutí do jednoho bodu tvoří prvek OLAP kostky. Na základě technologie OLAP jsou vytvářeny OLAP databáze, které představují jednu nebo více souvisejících OLAP kostek“.

2.3 Vrstvy, typy aplikací a hlavní komponenty BI

Za celou dobu vývoje se ustálila obecná koncepce vývoje řešení Business Intelligence. Jelikož rozmanitost problémů řešených pomocí nástrojů BI, stejně jako rozmanitost nástrojů samotných, vede k tomu, že tato obecná architektura má několik vývojových větví a její konkrétní aplikace v reálných situacích se podstatně liší.

Obecná koncepce (Novotný, 2005):

Vrstva pro extrakci, transformaci, čištění a nahrávání dat

Tato vrstva komponentů datové transformace pokrývá oblast sběru/přenosu dat ze zdrojových systémů do vrstvy pro ukládání dat v řešení BI:

- ETL systémy – neboli systémy pro extrakci, transformaci a přenos dat.
- EAI systémy – neboli systémy pro integraci aplikací.

Vrstva pro ukládání dat

Vrstva datových komponentů zajišťuje procesy ukládání, aktualizace a správy dat pro BI řešení:

- Datové sklady (Data Warehouse) – základní databázová komponenta řešení BI.
- Datová tržiště (Datamart) – subjektivě orientované analytické databáze, součást nebo nadstavba datového skladu.
- Operativní datová tržiště (Operational Data Store) – podpůrné analytické databáze.
- Dočasná úložiště dat (Data Staging Areas) – databáze pro dočasné uložení dat před jejich vlastním zpracováním do databázových komponent řešení BI.

Vrstva pro analýzy dat

Tato vrstva analytických komponentů pokrývá činnosti spojené s vlastním zpřístupněním dat a analýzou dat:

- Reporting – analytická vrstva, zaměřená na standardní nebo ad hoc dotazovací proces do databázových komponent řešení BI.
- Systémy On-Line Analytical Processing (OLAP) – vrstva zaměřená na pokročilé a dynamické analytické úlohy.
- Dolování dat (Data Mining) – systémy zaměřené na sofistikovanou analýzu velkého množství dat.

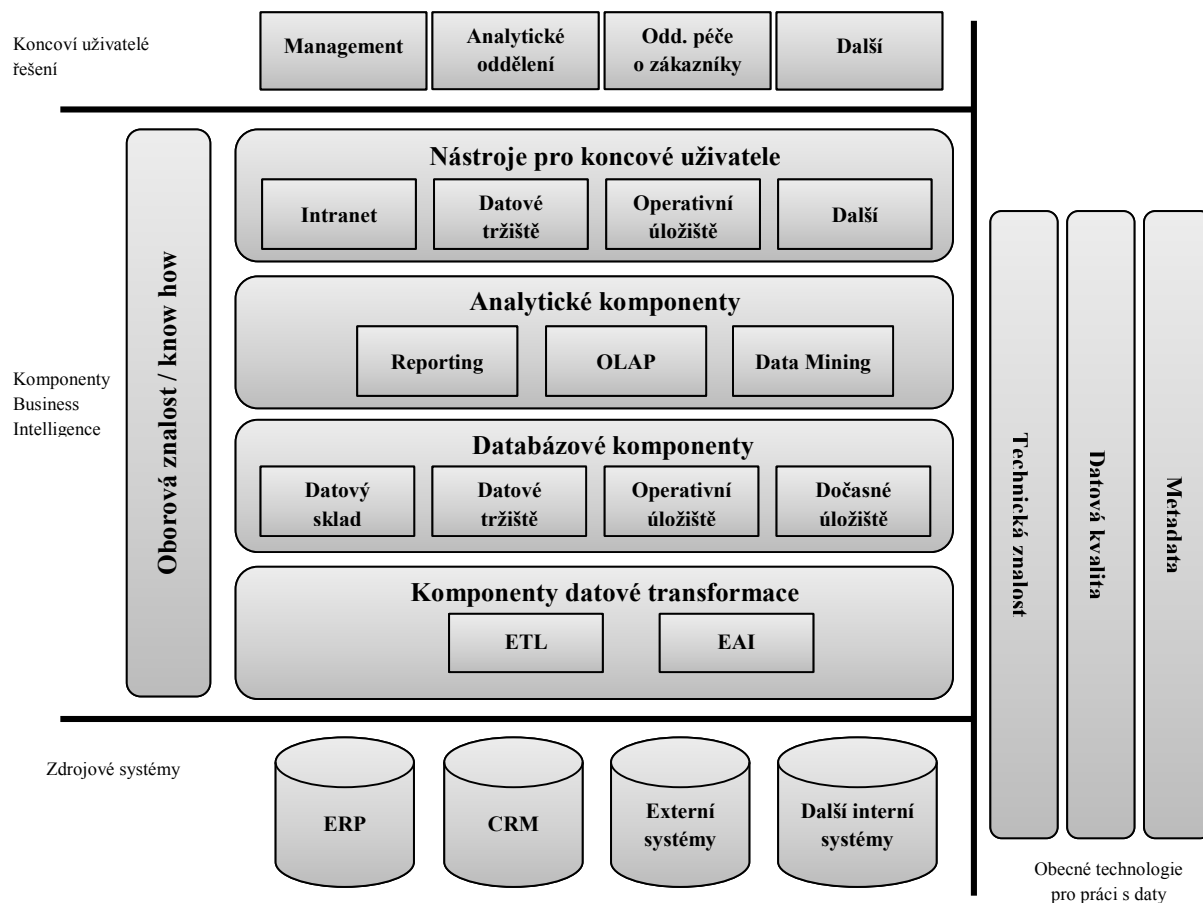
Prezentační vrstva

Představuje nástroje pro koncové uživatele, které zajišťují komunikaci koncových uživatelů s ostatními komponentami řešení BI, tedy zejména sběr požadavků na analytické operace a následnou prezentaci výsledků:

- Portálové aplikace založené na technologiích www.
- Systémy EIS – Executive Information Systems.
- Různé analytické aplikace.

Vrstva oborové znalosti

Zahrnuje oborovou znalost a tzv. best-practices nasazování řešení BI pro konkrétní situaci v organizaci.



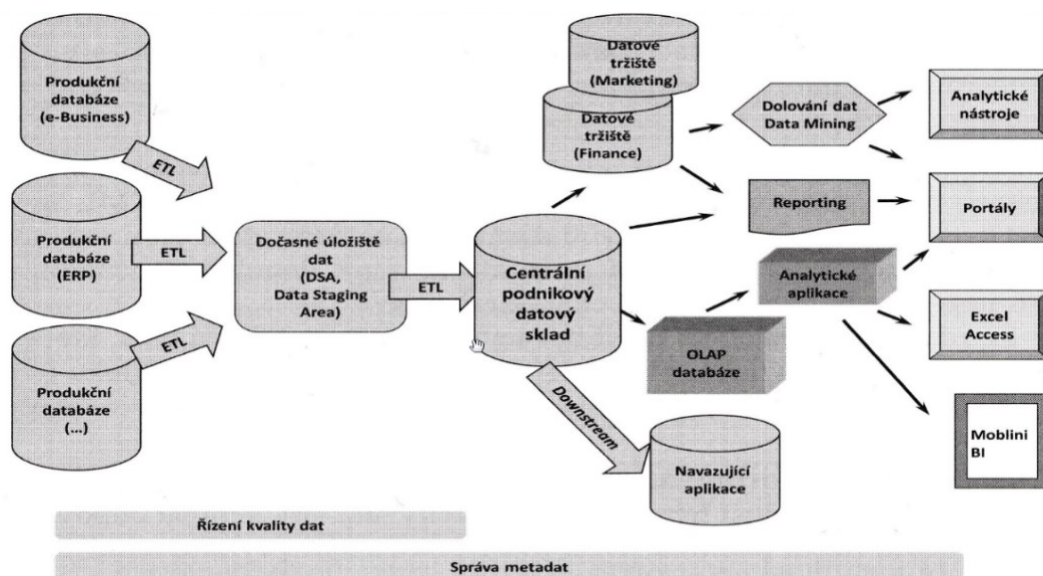
Obrázek 2-5 Obecná architektura řešení podle Novotného (2005, str. 27)

Aplikace BI kromě toho využívají níže uvedené obecné komponenty pro správu a manipulaci s daty:

- nástroje pro zajištění datové kvality, tedy nástroje starající se o to, aby data odpovídala realitě;
- nástroje pro správu metadat zabývající se popisem a dokumentací systému i probíhajících procesů;
- technickou znalost zahrnující programovací a technologicky závislé schopnosti implementačního týmu.

2.4 Hlavní komponenty řešení BI

Podkapitoly této části se budou zabývat jednotlivými komponentami komplexních řešení Business Intelligence jejíž složení se může výrazně měnit podle situace a potřeb daného podniku. To znamená od těch nejjednodušších řešení až po nejkomplexnější a také technologicky, finančně i pracovně nejnáročnější. Na Obrázku 2-6 je vidět obecné uspořádání jednotlivých komponent BI, které samozřejmě může v praxi nabývat různých variant a technologických realizací (Pour, a další, 2012).



Obrázek 2-6 Obecná architektura BI řešení (Pour, a další, 2012)

2.4.1 Produkční (zdrojové) databáze

Tyto databáze jsou většinou transakčního charakteru, ze kterých aplikace BI získávají data. Typickým příkladem to mohou být databáze aplikací ERP (Enterprise Resource Planning), SCM (Supply Chain Management) a CRM (Customer Relationship Management) specializující se na celé spektrum činnosti podniku od personálního oddělení přes finanční oddělení, účetnictví až po výrobu například. Produkční databáze jsou realizovány v nejrůznějších databázových systémech (MS SQL, ORACLE, Access, DB/2 atd.). Zdrojem dat pro BI mohou být i tabulkové soubory (Excel) nebo soubory v textovém vyjádření s oddělovači nebo s pevnou strukturou vět (flat files). Produkční systémy jsou hlavním a většinou jediným vstupem do BI. Zdrojem nemusí být pouze vnitropodnikové systémy, ale i externí jako je například databáze podnikatelských subjektů, výstupy statistického úřadu či vládních institucí apod. (Pour, a další, 2012).

Spektrum zdrojových databází pro BI je v praxi většinou velmi různorodé a heterogenní jak obsahově, tak technologicky. Smyslem řešení BI je zajistit analýzu těchto zdrojů z pohledu

potřeb řízení firmy, výběr relevantních dat pro řízení a následně jejich vzájemnou integraci. Tato část projektů BI je časově i finančně nejnáročnější, ale představuje nezbytný předpoklad pro úspěšně fungující aplikace BI (Novotný, 2005).

2.4.2 Extraction, Transformation and Loading – ETL

Jedna z nejvýznamnějších komponent celého řešení Business Intelligence je ETL, jinak řečeno datová pumpa. Úkolem datové pumpy je získat data ze zdrojových systémů (Extraction), poté je upravit do požadované formy a uspořádání (Transformation) a nakonec nahrát do datového skladu nebo tržiště (Load). ETL nástroje lze použít pro přenos mezi libovolnými databázemi nebo datovými soubory (textovými, tabulkovými, XML, atd.). ETL nástroje pracují v dávkovém režimu, to znamená, že data jsou přenášena v určitých časových intervalech, například denní nebo měsíční intervaly.

Podstatné charakteristiky ETL (Pour, Maryška a Novotný, 2012):

- ze zdrojových databází musí být vybrána pouze taková data, která jsou určena pro analytické, plánovací a rozhodovací aktivity podniku, což je jedním z prvních úkolů analytiků BI;
- data jsou transformována do nových datových struktur analytických databází, které musí být předem navrženy tak, aby nejlépe odpovídaly potřebám řízení podniku;
- data do BI vstupují z různých zdrojových databází (ERP, CRM, atd.), přičemž v těchto různých zdrojích mohou být tatáž data uložena vícekrát (různé databáze prodejců, zákazníků apod.). Do analytických databází však musí vstoupit pouze jednou. V transformační vrstvě, s využitím ETL, musí dojít ke konsolidaci dat, což představuje určení vstupujících dat s vyloučením duplicit a multiplicit;
- se zajištěním konsolidace dat úzce souvisí i dosažení potřebné kvality dat, to znamená vyloučení chyb, nepřesností atd.

2.4.3 Architektura datových úložišť

Systémy Business Intelligence obsahují různé typy datových úložišť. O jednotlivých typech je psáno v následujících odstavcích.

Datový sklad

Infrastruktura BI aplikací bývá založena na konceptu datového skladu (DS, Data Warehouse, DWH), jenž v dnešní době představuje nezbytnou součást a předpoklad

pro fungování nástrojů BI. Datový sklad představuje ucelenou databázi, optimalizovanou pro dotazování a analýzu dat, společně s nástroji, které umožňují dotazy, analýzy a kvalitní prezentaci výstupu. V DS jsou data integrována a ukládána jak z interních, tak i z externích zdrojů. Výsledkem je poskytnout čitelné, organizované a v reálném čase dostupné informace z maxima podnikových databází i externích zdrojů, které jsou ve velkém rozsahu využitelné při řízení podniku. Datový sklad je dlouhodobým úložištěm, kde data přibývají po jednotlivých dávkách.

V datovém skladu platí (Tvrdíková, 2008):

- odezvy na dotaz nemusejí být okamžité;
- připouští se určitá redundance dat;
- data z datového skladu se nikdy nelikvidují (mohou se provádět agregace údajů a zálohy na externí média);
- zdroje datového skladu mohou být uloženy ve zcela odlišných strukturách, formátech, mohou být uloženy na různých médiích atd.

Datové sklady fungují na principu již zmiňovaného ETL a využívají ETL nástroje. Prostředí DS musí poskytovat dostatek nástrojů a metod pro zvládnutí ETL. Klíčovou fází je validace a čištění, protože základní podmínkou využitelnosti datového skladu je to, že obsahuje důvěryhodná a správná data. Jakékoliv metody na využití dat, která nepracují se správnými daty, jsou podstatě bezcenné. Pro integrovaný centrální datový sklad je požadavek konzistence naprosto zásadní (pouze jedna verze pravdy) a jsou v současnosti považovány za nejlepší řešení (Tvrdíková, 2008).

Nevýhody, výhody a vlastnosti datového skladu (Tvrdíková, 2008):

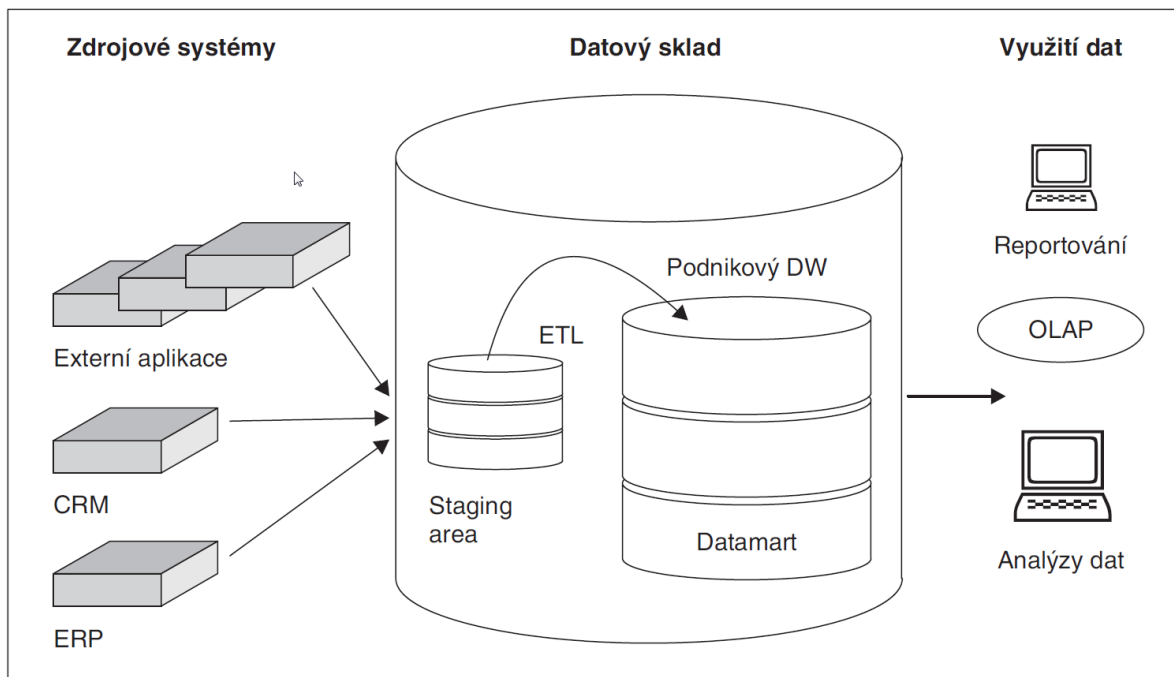
Nevýhody:

- složitější realizace;
- pomalejší implementace;
- sekundární načítací procesy (z centrálního DS do datového tržiště).

Výhody:

- konzistentní obsah DS;
- menší počet načítacích procesů z provozních systémů (primární načítací procesy);

- jednodušší správa načítacích procesů;
- snazší vytváření nových datamartů (detailní data jsou k dispozici již v DS).



Obrázek 2-7 Schéma datového skladu (Tvrdíková, 2008, str. 109)

Vlastnosti:

- integruje data z různých zdrojů do jednoho systému;
- obsahuje historii – k dispozici jsou data i za několik let zpátky;
- data jsou uložena na různých úrovních sumarizace;
- data se periodicky načítají z provozních systémů (nejlépe v noci nebo o víkendu);
- uživatelé data pouze čtou (nezařádají nová data ani nemění stávající);
- data jsou uspořádána podle jednotlivých subjektů;
- data jsou užívána pomocí různých aplikací pro prezentace a analýzy.

Jeden ze zakladatelů data warehousingu William Inmon (2005) definuje datový sklad jako: „integrovaný, konsolidovaný, subjektově orientovaný, stálý a časově rozlišený souhrn dat, uspořádaný pro podporu potřeb managementu“.

Význam jednotlivých pojmů:

- integrovaný - data jsou ukládána v rámci celého podniku a ne pouze v rámci oddělení;

- konsolidovaný - data jsou konsolidována z různých zdrojů, struktur a forem do jedné výsledné formy, která představuje jednu verzi pravdy;
- subjektivě orientovaný - data jsou rozdělována podle jejich typu, ne podle aplikací, ve kterých vznikla;
- stálý – datové sklady jsou navrhovány převážně pouze pro čtení, do datových skladů nelze uživatelskými nástroji data dodávat nebo je měnit a jsou v DS po celou dobu existence;
- časově rozlišitelný – načítaná data musí sebou nést i dimenzi času (historii dat), aby bylo možné provádět analýzy za určitá období.

Nezávislé datové tržiště

Datová tržiště (data marts) jsou koncipována na obdobném principu jako datové sklady, určena jsou však pro vymezený okruh uživatelů (oddělení, divize, pobočka atd.). Jsou to vlastně decentralizované datové sklady postupně se integrující do komplexního datového skladu. Nezávislá datová tržiště jsou samostatná datová úložiště pro jednotlivé aplikace nebo útvary a jsou to problémově orientované datové sklady umožňující provádět „ad hoc“ analýzy. Níže jsou uvedeny výhody a nevýhody (Tvrdíková, 2008).

Nevýhodou datových tržišť je:

- možná nekonzistence mezi jednotlivými data marts;
- komplikované načítací procesy (jejich velký počet, náročné na údržbu);

Výhodou je:

- snazší a rychlejší implementace;
- uživatelé rychleji získávají přínosy z investic do datového skladu;
- snižuje se riziko při zavádění datového skladu.

Dočasné úložiště dat

Důvodem existence dočasného úložiště dat (Data Staging Area, DSA) je dočasné uložení extrahovaných dat z produkčních databází s cílem zajistit jejich přípravu a potřebnou kvalitu před vstupem do datového skladu. Data v dočasném úložišti jsou detailní, neagregovaná, často nekonzistentní a bez časové dimenze. DSA obsahuje pouze aktuální data,

to znamená že, po jejich zpracování v DSA a přenosu do datového skladu nebo tržiště, se z DSA odstraní.

DSA napomáhají (Tvrdíkové, 2008):

- minimalizovat dopad transferu dat na výkonnost vytížených zdrojových systémů;
- u systémů pracujících například s textovými soubory (tam, kde je potřeba konvertovat data do databázového formátu);

a obsahují data, která jsou:

- detailní – nejsou agregovaná;
- nekonzistentní – nejsou kontrolována proti ostatním datům v DS;
- pouze aktuální – neobsahují historii (pouze aktuální data ze zdrojových systémů);
- měnící se – při každém výběru se berou pouze data, která ještě nebyla zpracována, po přenosu do dalších komponent řešení BI, se data z DSA vymažou;
- ve stejné struktuře, v jaké jsou uložena ve zdrojových systémech.

Operativní úložiště dat

Operativní úložiště dat (ODS) se nenachází ve všech řešeních BI. Existují dva základní přístupy k definici ODS.

První definuje ODS jako jednotné místo datové integrace aktuálních dat z primárních databází. Jedná se o zdroj pro sledování konsolidovaných agregovaných dat s minimální odezvou po zpracování (téměř v reálném čase). Většinou slouží jako centrální databáze základních číselníků nebo pro podporu interaktivní komunikace se zákazníkem (například call-centrum, kdy ODS dodává aktuální data o zákazníkovi). Tyto databáze podporují vkládání a modifikaci dat v reálném čase a jsou typicky napojeny na EAI platformy.

Druhý přístup definuje ODS jako databázi navrženou s cílem podporovat relativně jednoduché dotazy nad malým množstvím aktuálních analytických dat. Na rozdíl od prvního přístupu obsahuje ODS jen aktuální záznamy vybraného množství dat.

Stejně jako DSA, i ODS obsahuje data bez historie (aktuální) měnící se po každém nahrání. Oproti DSA však ODS obsahuje data konsolidovaná, konzistentní, subjektivě orientovaná a v některých případech i agregovaná (Novotný, 2005).

Zásadní rozdíl mezi ODS a DSA je v tom, že zatímco ODS jsou budována za účelem zpřístupnit uživatelům nebo ostatním systémům data pro analýzy nebo dotazy s minimálním zpožděním od jejich vzniku, do DSA nemají koncoví uživatelé přístup (Tvrdíková, 2008).

2.4.4 OLAP databáze

OLAP databázím se už věnovala kapitola 2.2.3. Pro připomenutí OLAP databáze představují jednu nebo několik souvisejících OLAP kostek, které, na rozdíl od datových skladů, obsahují předzpracované agregace dat podle definovaných hierarchických struktur dimenzí a jejich kombinací (Novotný, 2005).

2.4.5 Reporting

Reportingem se rozumí činnosti spojené s dotazováním se do databází pomocí standardních rozhraní těchto databází, například SQL dotazy. Tímto tématem se bude zabývat kapitola 2.7 (Pour, a další, 2012).

2.4.6 Analytické aplikace

Analytické aplikace jsou typem aplikací BI, pro které platí, že (Pour, Maryška a Novotný, 2012):

- jsou navrhovány pro poskytování manažerských informací, umožňují sledovat firemní procesy, plnění cílů organizace a podobně;
- poskytují nástroje pro online analýzy zahrnující analýzy trendů, dril up, dril down, slice and dice a identifikaci výjimek a podobně;
- jsou jednoduše ovladatelné a zajišťují vysokou vypovídací hodnotu prostřednictvím grafického uživatelského prostředí.

Analytické aplikace se mohou provozovat v různém technologickém prostředí. Většinou jsou realizovány aplikacemi nad OLAP databázemi, protože je požadována flexibilita. Tyto aplikace se vytvářejí a provozují pomocí různých specializovaných produktů (Talend studio, Microsoft SQL Server atd.) nebo kancelářských prostředků (Excel, Access atd.) (Pour, a další, 2012).

2.4.7 Dolování dat

Dolování dat, anglicky data mining, je proces extrakce relevantních předem neznámých nebo nedefinovaných informací z rozsáhlých databází. Jedná se o analýzy odvozené z obsahu, nikoli předem specifikované uživatelem. Taktéž se jedná o odvozování prediktivních informací,

ne pouze deskriptivních. Pour, Maryška a Novotný (2012, str. 27) píší: „dolování dat slouží manažerům především k objevování nových skutečností, umožňují testovat hypotézy, odhalují skryté korelace mezi ekonomickými veličinami a podobně“.

Dolování dat je založeno na množství matematických a statistických technik. Příkladem jsou rozhodovací stromy, což je prediktivní model zobrazující data v podobě stromu, kde každý kořen určuje kritérium pro následné rozdělení dat do jednotlivých listů (Rud, 2001).

2.4.8 Nástroje pro řízení kvality dat a správu metadat

S růstem nasazení analytických aplikací, zažívají nástroje pro zajištění datové kvality prudký rozvoj. Tyto práce však musí probíhat nad korektními daty, dokumentující reálnou situaci podniku. Nástroje pro zajištění datové kvality se zabývají zpracováním dat s cílem zajistit jejich požadované vlastnosti.

Nástroje pro správu metadat zase získaly na větší důležitosti až s rozvojem business intelligence. Metadata jsou definována jako data o datech a slouží pro dokumentaci konkrétních implementací informačních systémů podniku. Metadata popisují veškeré informační systémy i jejich jednotlivé části. Metadata z pohledu BI zahrnují zejména datové modely, popisy funkcí, business a transformačních pravidel, reportů či požadavků na reporty a podobně. Podstatnou součástí metadat jsou i vymezení a interpretace ekonomického, respektive podnikového obsahu a podstaty jednotlivých analytických aplikací a reportů, jejich součástí, jejich způsobu vzniku, pravidel vytváření atd. Pokud by se to dalo shrnout do jedné věty tak metadata definují business obsah zpracovávaných a prezentovaných dat (Pour, a další, 2012).

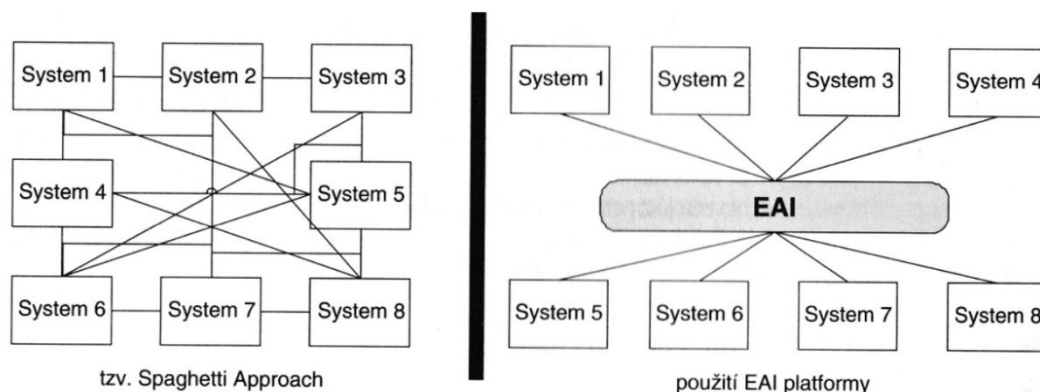
2.4.9 Enterprise Application Integration – EAI

Nástroje EAI vznikly a jsou v drtivé většině využívány ve vrstvě zdrojových systémů. Cílem EAI je integrovat primární podnikové systémy a redukovat počet jejich vzájemných rozhraní. EAI pracují ve dvou úrovních (Pour, a další, 2012):

- na úrovni datové integrace, kde jsou EAI platformy využity pro integraci a distribuci dat;
- na úrovni aplikační integrace, kde jsou EAI platformy využity nejen pro integraci a distribuci dat, ale především pro sdílení určitých vybraných funkcí informačních systémů.

Uplatnění ETL nástrojů se nachází zejména ve vrstvě datové integrace, kde jsou využity pro přenos dat do datových úložišť v reálném čase. Doplnuje tak dávkový přenos (ETL)

a umožňuje vznik datových skladů pracujících v reálném čase, tzv. Real-Time Data Warehouse. Na Obrázku 2-8 lze vidět smysl použití EAI platformy (Pour, a další, 2012).



Obrázek 2-8 Požití EAI platformy (Novotný, 2005, str. 30)

2.5 Úvod do analytických služeb

Pro lepší pochopení problematiky multidimenzionality je potřeba rozlišovat mezi klasickou relační a multidimenzionální analytickou databází.

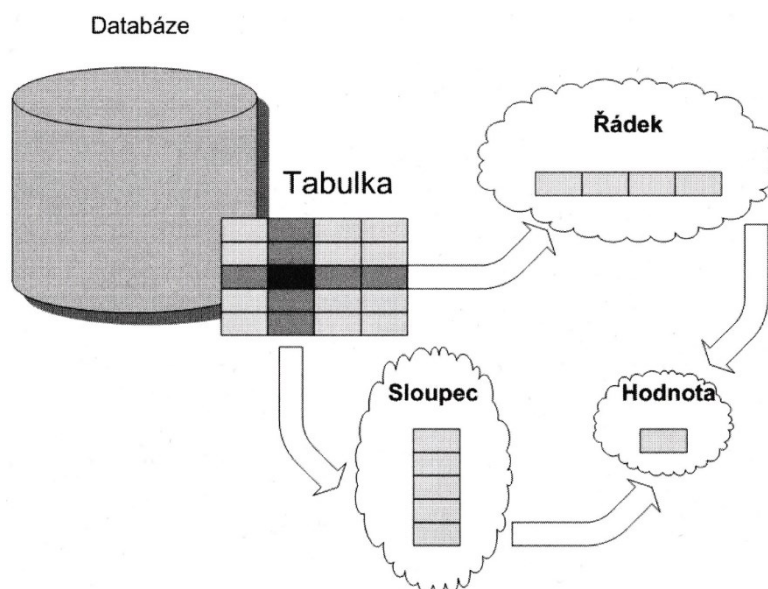
2.5.1 Relační databáze

V relačních databázích, které jsou nasazeny v nejrůznějších oblastech (banky, supermarkety atd.), probíhá velké množství transakcí v reálném čase, nebo dochází ke změnám údajů při monitorování nějakého procesu. Takové databáze se označují OLTP (On-Line Transaction Processing) databáze nebo taktéž operační databáze, protože se v nich vykonává velké množství operací. Transakční databáze, do nichž se ukládají aktuální operační data, jsou organizována jako relační, což znamená, že jsou uložena v databázových tabulkách, mezi kterými jsou relační vztahy vyplývající z aplikační logiky. Definice databáze podle E. F. Codd zní: „databáze je úložiště dat, která jsou uložena a zpracovávána nezávisle na aplikačních programech. Databáze zapouzdřují nejen vlastní data, ale i relační vztahy mezi jednotlivými prvky a objekty v databázi, schémata popisující struktury dat a integritní omezení“.

Databázová tabulka

V prostředí relačních databází jsou data, v nichž se skrývají potenciační informace, uložena v klasických databázových tabulkách skládající se z hodnot umístěných na průsečíku řádků a sloupců, kdy každý řádek v tabulce je zpravidla jednoznačně určený nějakým identifikátorem (Lacko, 2009). Průkopník v oblasti teorie relačních databází Dr. E. F. Codd definoval tři podmínky minimální relačnosti:

- všechna data v databázi jsou uložena v tabulkách;
- fyzická struktura dat a jejich uložení je nezávislé a od uživatele zcela odstíněné;
- pro práci s daty v databázi předpokládáme existenci databázového jazyka umožňujícího realizovat minimálně operace selekce, restrikce, projekce a spojení.



Obrázek 2-9 Princip relační databáze (Lacko, 2009, str. 167)

Relační vztahy

Aby vůbec bylo možné aplikovat relační teorii, je nutné vytvořit více databázových tabulek a definovat relační vztahy mezi nimi. Primární klíč je jedinečný identifikátor určující jedinečnost každého záznamu (řádku) tabulky. Může to být sloupec, případně kombinace více sloupců. Hodnota pole primárního klíče musí být v rámci tabulky jedinečná. Pole klíče musí obsahovat konkrétní hodnotu, což znamená, že nesmí nabývat hodnoty null. Bez primárního klíče není možné definovat relace mezi tabulkami. Cizí klíč je pak sloupec, případně kombinace sloupců, které jsou propojené s primárním klíčem v jiné tabulce.

Vztahy mezi entitami

Relace mezi tabulkami popisují vztahy mezi objekty reálného světa, které tyto tabulky představují. Entity můžeme rozdělit na primární a sekundární, mezi kterými je vztah neboli relace. Uvádí se tyto typy relací (Lacko, 2009, str. 169):

- Vztah 1:1 (one-to-one) – první entitě, tedy záznamu v databázové tabulce, odpovídá maximálně jedna entita, záznam z jiné tabulky. Což v překladu znamená, že každý řádek primární tabulky lze svázat právě s jedním řádkem sekundární tabulky.

- Vztah 1:N (one-to-many) – každý řádek primární tabulky lze svázat s jedním nebo více řádky sekundární tabulky. Při tomto vztahu je důležitý směr. Například jeden zákazník může mít více objednávek (ne naopak).
- Vztah N:M (many-to-many) – více řádků primární tabulky je svázáno s více řádky tabulky sekundární. Jelikož většina databázových systémů nedokáže pracovat se vztahem N:M, tak se tento vztah implementuje pomocí spojovací tabulky. Vztah N:M se dá rozdělit na dva vztahy typu N:1.

Normalizace databází

Lacko (2009, str. 169) říká, že: „při návrhu relačního schématu je nutné dodržovat určitá pravidla, která se nazývají reálné formy. Normalizace obvykle vede k odstranění redundancí a značně zefektivňuje práci s databázovými tabulkami, pro něž platí, že v čím vyšších normálních formách databázové tabulky jsou, tím lépe by se s nimi z hlediska aplikační logiky mělo pracovat“. Dále Lacko (2009, str. 170) popisuje tyto normální formy:

- První normální forma (1NF) – tabulka splňuje první normální formy tehdy, když jsou všechny atributy (sloupce) atomické, to znamená dále nedělitelné. Například jméno Adam Rybář nebude v jednom sloupci (celé jméno), ale ve dvou sloupcích s názvem jméno a příjmení.
- Druhá normální forma (2NF) – tabulka splňuje podmínku pro zařazení do druhé formy tehdy, když splňuje podmínku 1NF a každý atribut kromě primárního klíče musí být úplně závislý na celém primárním klíči. Tato forma se týká tabulek, které mají více primárních klíčů. V praxi to znamená, že když máme tabulku s atributy zboží, dodavatel, email dodavatele a cena s primárními klíči na prvních dvou attributech, tak email dodavatele nezávisí na zboží, ale pouze na dodavateli, tím pádem email nemá žádný vliv na cenu, která je tvořena pouze kombinací zboží a dodavatele. Řešením je rozdělit původní tabulku na více tabulek, kdy jedna bude obsahovat atributy zboží, dodavatel a cena, a druhá tabulka dodavatel a email dodavatele.
- Třetí normální forma (3NF) – tabulka je ve třetí normální formě, když je ve druhé normální formě a zároveň neexistují závislosti neklíčových sloupců tabulky. Otázkou je pak jestli například vytvářet tabulku zvlášť pro PSČ s několika tisíci záznamy, což by ovlivnilo výkon databáze, nebo nechat tabulku v 2NF a pro PSČ vytvořit další atribut, i když se už v tabulce nachází atribut s názvem město.

Normalizované tabulky umožňují (Lacko, 2009, str. 172):

- velký počet paralelních přístupů;
- rychlé a operativní změny dat v tabulkách;
- využití indexů, transakcí;
- optimalizaci struktury dat s využitím normalizovaných tabulek.

2.5.2 Multidimenzionální databáze

„Databáze s multidimenzionálními strukturami dat slouží jako podklad pro získání sumarizovaných a agregovaných informací. Do multidimenzionálních databází a datových skladů se ukládají upravená a vyčištěná data. Oproti relačním databázím se používají převážně nenormalizované tabulky, které je možné rozdělit na dva druhy, a to na tabulky faktů a tabulky dimenzí“ (Lacko, 2009, str. 172).

K hlavním výhodám multidimenzionální databáze patří (Lacko, 2009, str. 172):

- rychlý komplexní přístup k velkému objemu dat;
- přístup k multidimenzionálním a relačním datovým strukturám;
- možnost komplexních analýz;
- silné schopnosti pro modelování a prognózy.

Nevýhodou můžou být vyšší nároky na kapacitu úložiště, problémy při změně dimenzí, bez přizpůsobení časové dimenze a podobně. Analytické databáze jsou často označovány pojmem OLAP (OnLine Analytical Processing) zahrnující struktury data analytické služby, které slouží pro analýzu velkého množství dat.

Multidimenzionální databázový model

Lacko (2009, str. 172) uvádí: „většina dat je organizována v relačních tabulkách, kde každý řádek takové tabulky se vztahuje k nějakému předmětu, události, nebo jejich části. Výsledkem agregace a analýzy dat je obvykle multidimenzionální datová struktura – kostka. Zjednodušeně by se dalo říct, že kostka je ekvivalentem tabulky v relační databázi. OLAP systémy se typicky využívají pro analýzu velkého množství dat“. Výsledkem analýzy jsou souhrny a reporty, které slouží manažerům jako podklady pro rozhodování v oblasti řízení firmy nebo ekonomických a technologických procesů a podobně. Pro výpočet OLAP kostek je nutné vykonat velké množství výpočtů a agregací téměř v reálném čase. Každá kostka

má několik dimenzí, například čas, region, produkt. Data jsou pak logicky situována na průnicích jednotlivých dimenzí. Fyzicky jsou data uložena v tabulkách relační databáze a multidimenzionální struktura je dosažena relačními vazbami mezi tabulkami faktů a tabulkami dimenzí (Lacko, 2009).

2.5.3 Porovnání relačního a multidimenzionálního modelu

Výhody a nevýhody relačních databází (Lacko, 2009, str. 174):

výhody:

- potenciál odborníků ve firmách, kteří tento model mnoho let rutinně používají;
- potenciál softwaru a vývojových nástrojů pro vývoj a ladění aplikací a pro generování reportů;
- použitelnost v transakčních databázích i datových skladech;

nevýhody:

- absence komplexních analytických nástrojů;
- potenciální omezení objemů dat, ke kterým je možné v rozumném čase přistoupit.

Výhody a nevýhody multidimenzionálních databází (Lacko, 2009, str. 175):

Výhody:

- rychlý komplexní přístup k velkému objemu dat;
- přístup k multidimenzionálním a relačním datovým strukturám;
- možnost komplexních analýz;
- silné schopnosti pro modelování a prognózy;

nevýhody:

- problémy při změně dimenzí bez přizpůsobení časové dimenze;
- vyšší nároky na kapacitu úložiště.

2.6 Problematika analytických databází

Termín OLAP zavedl Dr. E. F. Codd, podle něhož je to „volně definovaná řada principů, které poskytují dimenzionální rámec pro podporu rozhodování“.

OLAP by měl poskytovat uživateli multidimenzionální model odpovídající jeho podnikatelským potřebám tak, aby tento model mohl využívat pro analýzu shromážděných dat. Technologie systému OLAP by měla být pro uživatele transparentní, aby mohl svou produktivitu a odbornost využít v koncových nástrojích a prostředí, například v programu Microsoft Excel. OLAP systém by měl přistupovat jen k datům, které jsou potřebné pro analýzu. Uživatel by neměl, při růstu počtu záznamů v databázi, pocítit snížení výkonu. Systém OLAP by měl podporovat víceuživatelský přístup ke konkrétnímu modelu. Systém OLAP musí umět dokázat rozeznat dimenzionální hierarchie a automaticky provádět asociované kumulované kalkulace v rámci dimenzí, musí umět „dril down“ a „drill up“ a musí existovat schopnost uspořádat řádky, sloupce a buňky způsobem, který umožní analýzu intuitivní vizuální prezentací analytických sestav. Pro vytvoření OLAP struktury je potřeba dvou druhů dat - fakta a dimenze (Lacko, 2009).

2.6.1 Fakta

Tabulka faktů je zpravidla největší tabulka databáze a obsahuje velký objem dat. Prvotní fakta (například objem prodeje) se mohou kombinovat anebo vypočítat pomocí jiných faktů a vytvořit tak měrné jednotky, které lze uložit v tabulce faktů.

Některá měřítko jsou aditivní (je možné je sčítat i napříč časem), kdy příkladem může být obrát za rok, který je možné získat sečtením obrátů za jednotlivé měsíce.

Jiná měřítko lze kvantitativně sčítat, kdy zase například celkový objem zásob je dán součtem objemů zásob v jednotlivých skladech, ale nelze je sčítat napříč časem. Je to logické, protože například zásoby, které byly skladem v lednu, mohou zůstat skladem i další měsíc (Lacko, 2009).

2.6.2 Dimenze

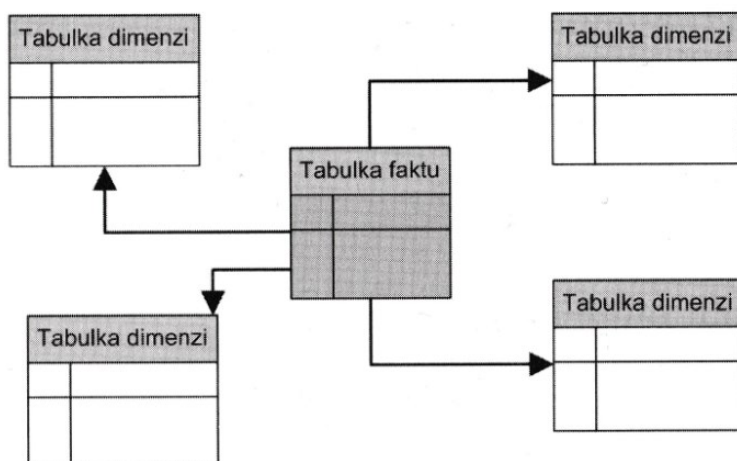
„Dimenze obsahují logicky nebo organizačně hierarchicky uspořádaná data. Jsou to vlastně textové popisy obchodování“ říká Lacko (2009, str. 177). Tabulky dimenzí jsou normálně menší než faktové tabulky a data se nemění tak často. Dimenze odpovídají na otázky proč a jak. Dimenze obecně obsahují relativně neměnná data, například geografické nebo časová dimenze. Není to však pravidlem, taková dimenze zákazníku nebo produktů se u větších podniků určitě aktualizuje častěji. Taková dimenze zákazníka může obsahovat atributy jako věk, pohlaví, bydliště atd.

Tabulky dimenzí obvykle obsahují hierarchickou stromovou strukturu. Typickým příkladem takové dimenze může být časová dimenze, kdy hierarchie může mít tyto úrovně: rok >> kvartál >> měsíc >> týden >> den (Lacko, 2009).

2.6.3 Schémata

„OLAP kostka se vytváří na základě dimenzionálního modelu, který má určité topologické uspořádání, kterému se jednoduše říká schéma“, píše Lacko (2009, str. 178). Mezi typická schémata patří hvězdicové schéma (star scheme) a schéma sněhové vločky (snowflake scheme).

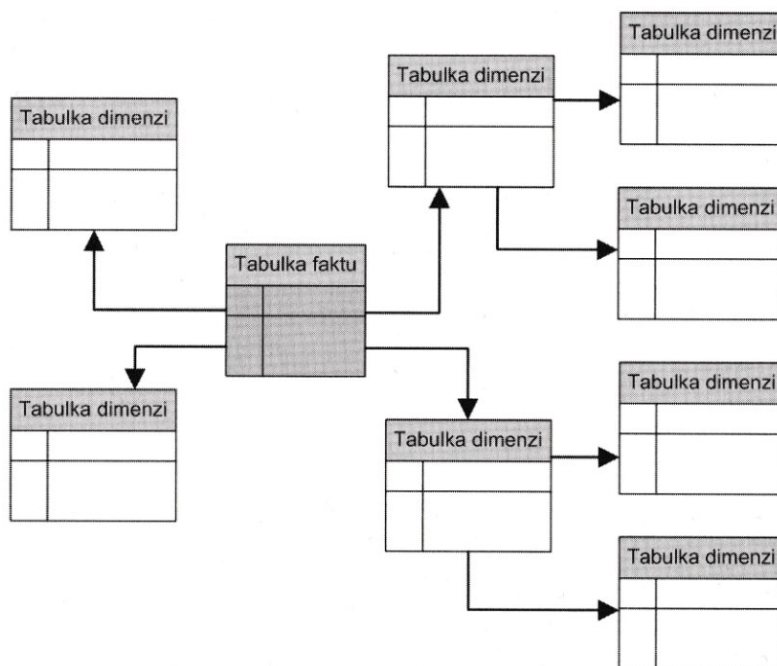
Hvězdicové schéma se skládá s tabulky faktů obsahující cizí klíče, které se vztahují k primárním klíčům v tabulkách dimenzí. Toto schéma nemá normalizované dimenze ani relační propojení mezi tabulkami dimenzí, proto je velmi snadno pochopitelné, ale v důsledku právě nenormalizovaných dimenzí je vytvoření tohoto modelu relativně pomalé. Typickým příkladem takové nenormalizované tabulky je časová dimenze, kdy z důvodu globalizace vzniká redundance dat v podobě uvádění dní, měsíců atd. v různých jazycích nebo uvádění dat v různých formátech (2015-25-1, 2015/25/1, 20152501,...). Všechna data potřebná pro časovou dimenzi se tak dají získat najednou a není nutné je skládat z relačních tabulek (Lacko, 2009).



Obrázek 2-10 Hvězdicové schéma (Lacko, 2009, str. 178)

Pour, Maryška a Novotný (2009, str. 19) píše, že: „v některých případech je vhodnější místo schématu hvězdy použít schéma sněhové vločky a to zejména tehdy, pokud se data v dimenzionálních tabulkách častěji aktualizují. V těchto případech se proto dimenzionální tabulky normalizují, to znamená, že se tabulka rozdělí podle hierarchických úrovní dimenze do více tabulek, aby se stejná data v záznamech neopakovala“. Lacko (2009, str. 179) pak doplňuje: „takto organizovaný model dimenzí umožňuje rychlejší zavedení dat

do normalizovaných tabulek, ale v porovnání s hvězdicovým schématem má nižší dotazovací výkon, neboť obsahuje větší množství spojení tabulek“.



Obrázek 2-11 Schéma sněhové vločky (Lacko, 2009, str. 180)

Databáze datového skladu či datových tržišť nemůže být řešena pouze jedním schématem typu hvězda nebo sněhové vločky, ale na základě značného množství jejich kombinací. Základní principem je, že se sledované ukazatele dělí do tabulek faktů podle logické příbuznosti a tedy i podle stejných přiřazených dimenzí a stejných složených klíčů. Takto se vytvářejí jednotlivé schémata, která se pak provazují pomocí společných, respektive sdílených dimenzí (Pour, a další, 2012).

2.6.4 Úložiště multidimenzionálních dat

Lacko (2009, str. 180) uvádí následující varianty technologie OLAP.

- MOLAP (Multidimenzionální OLAP) – pro MOLAP se data získávají buď z datového skladu, nebo z operačních zdrojů. Mechanismus OLAP poté uloží analytická data ve vlastních databázových strukturách a sumářích. Data v úložišti typu MOLAP se pak budou ukládat jako předem vypočítané pole. Databáze je organizovaná tak, aby umožnila rychlé získávání příslušných dat z více dimenzí. Hlavní výhodou MOLAP je maximální výkon vzhledem k dotazům uživatele, nevýhodou pak je redundance dat, protože jsou uložena jak v relační databázi, tak v databázi multidimenzionální.

- ROLAP (Relační databázový OLAP) – získává data pro analýzy z relačního datového skladu. Tato data z relačních databází se po zpracování předkládají uživateli jako multidimenzionální pohled. Data i metadata se v případě ROLAP ukládají jako záznamy v relační databázi. OLAP server pak dynamicky používá metadata pro generování SQL příkazů potřebných k získávání dat požadovaných uživatelem. Data zůstávají uložena v relačních databázích, takže nevzniká problém s redundancí dat.
- HOLAP (Hybridní OLAP) – je kombinací předchozích přístupů využívající výhod obou přechozích typů a tím pádem se částečně eliminují jejich nevýhody. Data jsou uložena v relačních databázích a napočítané agregace se ukládají do multidimenzionálních struktur. Při dotazování se data ukládají do multidimenzionální paměti cache. V tomto hybridním řešení relační databáze ukládá množství detailních dat a multidimenzionální model ukládá sumární data.

Novotný (2005, str. 33) pak doplňuje jednu variantu navíc:

- DOLAP (Desktop OLAP) – je nejmladší architektura OLAP databází a umožňuje připojit se k centrálnímu úložišti OLAP dat a stáhnout si potřebnou podmnožinu kostky na lokální počítač. Veškeré analytické operace jsou pak prováděny nad touto lokální kostkou, aniž by byl uživatel připojen k serveru. Výhodné je to zejména pro mobilní aplikace a obecně tento přístup slouží pro podporu mobilních uživatelů.

2.7 Reportovací služby

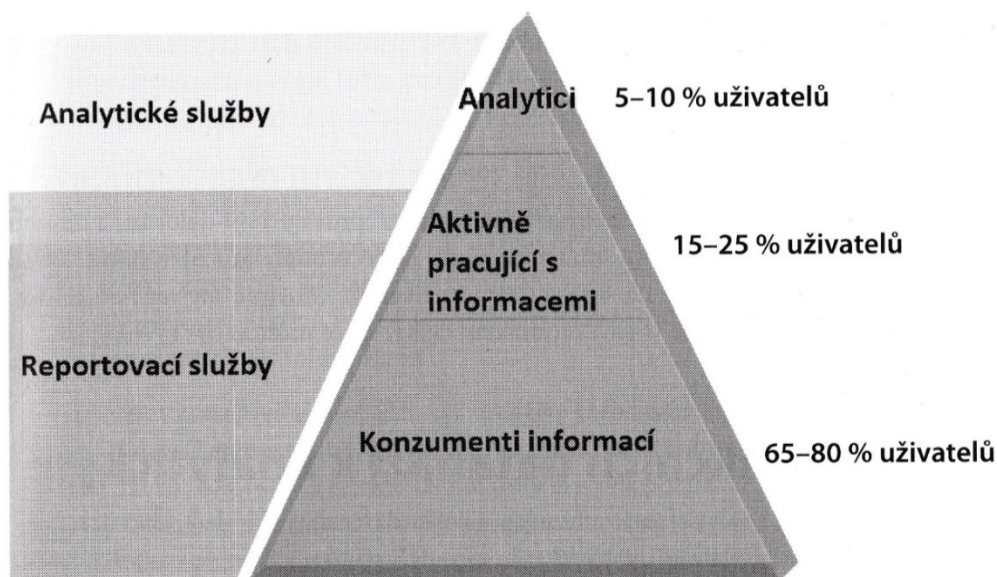
Jelikož je vytvoření dashboardu stěžejním cílem diplomové práce, je třeba reportům, kterými je dashboard tvořen, vymezit kapitolu zvlášť. V kapitole 2.8 pak bude věnována pozornost tomu, co je to dashboard a jakými prvky může být tvořen.

2.7.1 Úvod do reportingu

Na podnikové úrovni se generují různé druhy reportů, například pro finanční a obchodní oddělení. Data jsou buď v podnikové databázi, nebo v datových skladech. Výhodou je, že data jsou předzpracována a přetransformována a přenesena z produkčních systémů do datových skladů, případně datových tržišť. Reporty z reportovacích služeb potom vhodně doplňují data z analytických BI aplikací.

Na Obrázku 2-12 lze vidět, že pyramidové schéma využití jednotlivých BI technologií tvoří tři skupiny uživatelů:

- Analytici – jsou na vrcholu pomyslné pyramidy a podle průzkumů tvoří 5 až 10 % uživatelů BI. Jejich úlohou je vybírat data, která jsou vhodná pro analýzy, analyzovat je a na základě výsledků analýz poskytovat informace pro rozhodování. Analytici také dokážou na základě současného vývoje situace předpovídat trendy vývoje do budoucnosti.
- Aktivně pracující s informacemi – uživatelé z této skupiny, tvořící 15 až 25 %, data nejen analyzují a zpracovávají, ale taktéž je potřebují v různé formě zobrazovat.
- Konzumenti informací – tuto skupinu uživatelů tvoří 65 až 80 % uživatelů BI. Tito uživatelé dostávají výpisy dat ve formě dvojrozměrných sestav. Přístup k datům ve formě reportů pro tuto nejpočetnější skupinu zprostředkovávají reportovací služby.



Obrázek 2-12 Procentuální využití jednotlivých BI technologií (Lacko, 2009, str. 326)

Reportovací a analytické služby se částečně překrývají. Technicky lze říci, že obě služby si navzájem poskytují data. Reportovací služby umožňují generovat sestavy a reporty z relačních i analytických databází (Lacko, 2009).

2.7.2 Kategorizace reportů

„Úlohou reportingu je poskytovat ve vhodné formě a včas podklady pro podporu rozhodování na všech stupních organizační struktury“ uvádí Pour, Maryška a Novotný (2009, str. 133). Tyto podklady mohou být buď pouze statické, tedy pro čtení uživatelem nebo interaktivní s možností měnit formu a uspořádání reportů podle uvážení uživatele.

Reporty lze dělit několika způsoby. Pour, Maryška a Novotný (2012, str. 133) reporty dělí na standardní a ad-hoc.

- Standardní – se vytvářejí automaticky a pravidelně se distribuují. Mohou být statické nebo interaktivní (dynamické). Je možné měnit formu reportu, provádět operace, které zobrazují sledovaná fakta na různé úrovni a v různých kombinacích dimenzí, nebo je možné zobrazovaná data různě filtrovat. Standardní reporty by měli být dostupné pomocí různých koncových zařízení včetně mobilních (mobilní BI). Podmnožinou standardních reportů se chápou i dashboardy.
- Ad-hoc – tyto reporty si uživatelé vytvářejí sami bez podpory IT specialistů a pokrývají jejich aktuální požadavky, jenž stojí mimo standardní reporting. Nástroje pro tvorbu ad-hoc reportů umožňují efektivní tvorbu reportů s přístupem k různým zdrojům dat.

Další rozdělení, které nám nabízí Lacko (2009, str. 324), je podle filozofie a oblasti nasazení:

- Enterprise reporting – je reportování převážně na podnikové úrovni například pro obchodní, finanční oddělení atd. Výhodou je, že data jsou předzpracována a přetransformována v etapě ETL a přenesena z produkčních databází do datových skladů, případně datových trhů. Reporty pak vhodně doplňují data z analytických BI aplikací. Nasazení reportovacích služeb je převážně na úrovni podnikových portálů, k nimž uživatelé přistupují v rámci podnikového intranetu.
- Embedded – tyto reporty jsou integrální součástí aplikací.
- Business To Business (B2B) – generování reportů obchodním partnerům a zákazníkům. V těchto případech se občas zakrývají určité informace, které příslušný konzument reportu pro své rozhodování nepotřebuje. Jsou to zejména informace, které odhalují slabá místa firmy, které by mohl obchodní partner odhalit.

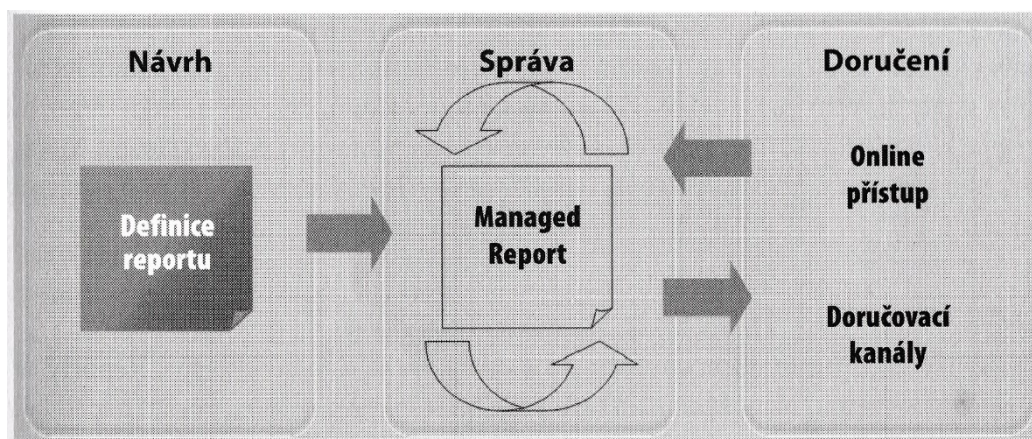
2.7.3 Životní cyklus reportu

Návrh reportu lze přirovnat například k vývoji aplikace. Životní cyklus reportu má tři fáze, kterými jsou (Lacko, 2009, str. 329):

- Návrh – report v etapě návrhu je jen jakási šablona určující k jakým datům a jakým způsobem se bude přistupovat. Data prezentována prostřednictvím reportů mohou

být zobrazována v nespočetných variacích tabulek, matic, grafů, nebo hypertextu. Variabilitu podporují parametry, filtry, seskupování dat a výpočet částečných výsledků pomocí agregačních funkcí. Hierarchicky uspořádaná data lze zobrazovat na různých úrovních granularity hierarchické struktury, od globální až po velmi detailní výsek množiny dat. Většina reportů je postavena na SQL dotazech, které zaručují další možnosti flexibility.

- **Správa** – v této etapě se reporty spravují a sladují se návrhy. Generování reportů může být buď aktuální, tzv. na požádání nebo na základě nějakých plánů či časových rozvrhů.
- **Doručení** – zde je důležitý způsob doručení a forma. Mohou se generovat přímo na požádání nebo na základě časových plánů a posílat například emailem. Formu si lze představit jako tabulky, grafy, obchodní grafiku a podobně. Lze si taktéž zvolit typ dokumentu nebo výstupní formát, do kterého se report vygeneruje. Mohou to být webové formáty (HTML, OWC), tisknutelné formáty (PDF, RTF, TIFF) nebo datové formáty (Excel, XML, CSV). Odběratel může reporty dostávat například ve formě emailů, případně pomocí sdílení souborů. Odběratele si také mohou reporty individuálně personalizovat.



Obrázek 2-13 Životní cyklus reportu (Lacko, 2009)

Na Obrázku 2-13 si lze všimnout, že mezi správou reportu a procesem jeho doručení probíhá oboustranná interakce (Lacko, 2009).

2.8 Dashboard

Dashboard je nástrojem pro vizualizaci podnikové výkonnosti sloužící vrcholovým pracovníkům jako účinná podpora pro rozhodování (Novotný, 2005).

Dashboardy jsou aplikace, které umožňují vizuální zobrazení nejdůležitějších informací, které jsou potřebné k dosažení určitého cíle nebo cílů, které jsou konsolidované a uspořádané na jedné obrazovce tak, že lze sledovat tyto klíčové informace na první pohled (Few, 2006).

Eckerson (2005) používá výraz Performance dashboard, což značí něco více než vizualizaci dat. Podle něj je to vícevrstvá aplikace postavená na základech BI datové integrační infrastruktury umožňující podnikům monitorovat, měřit a řídit výkonnost efektivněji.

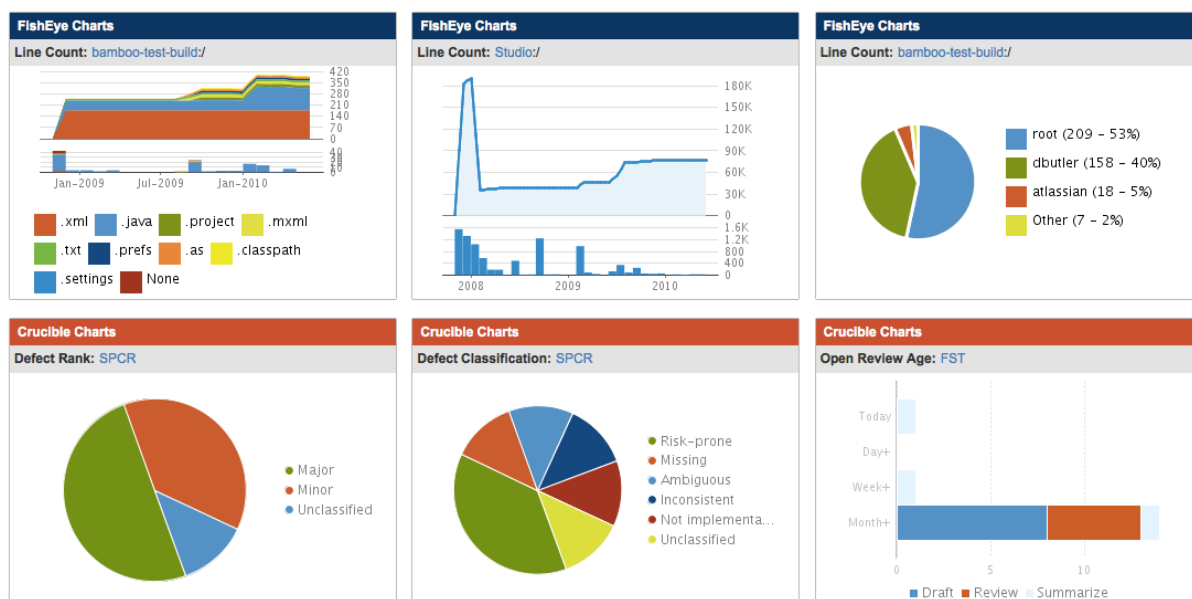
Dashboardy lze podle určení klasifikovat na strategické operační a analytické:

- Operační dashboardy slouží vedoucím pracovníkům a specialistům a jejich účelem je monitoring operací. Vztahuje se na oddělení nebo jednotlivce a obsahuje detailní informace s důrazem na monitoring. Aktualizace probíhají denně i častěji.
- Taktické dashboardy využívají vedoucí pracovníci na střední úrovni řízení a analytici pro měření postupu. Obsahují detailní i souhrnné informace určitého oddělení aktualizující se minimálně jednou týdně. Důraz je kladen na analýzu.
- Strategické dashboardy slouží pro vedoucí pracovníky nejvyšší úrovně řízení. Účelem je aplikovat strategii vztahující se na podnik a obsahující detailní i souhrnné informace. Aktualizují se podle potřeby měsíčně či čtvrtletně. Důraz je kladen na strategické řízení.

Dashboard musí splňovat tyto 4 vlastnosti (Few, 2006):

- Dashboard je vizuální zobrazení nějakých informací, kdy je vizualizace provedena graficky i textově, avšak s důrazem na grafickou formu. Grafická forma prezentace je totiž lépe a rychleji vnímána a zpracována lidským mozkem. Důraz při tvorbě dashboardu je tedy kladen na principy lidského vnímání informace.
- Dashboard zobrazuje informace potřebné pro dosažení jednoho či více cílů. Informace přitom nemusí spolu přímo souviset. Účelem dashboardu je vybrat pouze nejnutnější informace a pokusit se je logicky co nejvíce provázat a zjednodušit. V dashboardu by se neměly vyskytovat informace nesouvisící s aktuální činností.
- Dashboard by se měl vejít na jednu obrazovku. Důvod je ten, že jednotlivé informace získávají určitý význam až jako jednotný celek. Aby si člověk vůbec tento význam spojil, musí mít informace pokud možno všechny pohromadě před sebou na jediné stránce. Dashboard by měl právě tuto situaci uživateli co nejvíce ulehčit.

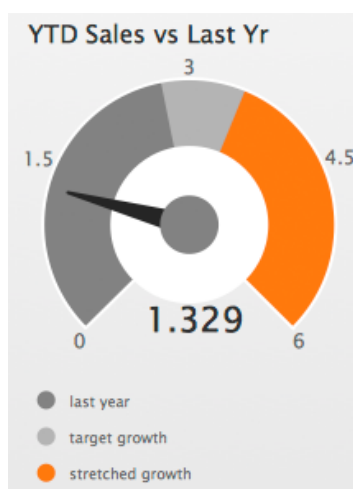
- Dashboard by měl podle Fewa být dostatečně intuitivní a přehledný, aby uživatel během prohlédnutí obrazovky rychle pochopil, čeho se dashboard týká a jaký je aktuální stav monitorované skutečnosti.



Obrázek 2-14 Ukázka dashboardu z webových stránek blogs.atlassian.com

Jak lze vidět na Obrázku 2-14, informace se na dashboardu dají zobrazovat několika způsoby. Mezi základní způsoby zobrazení patří grafy, tabulky, ikonky a v poslední době i mapy.

Grafy jsou základem většiny dashboardu a zobrazují kvantitativní data. K zobrazení se dají použít všechny typy grafu jako je například sloupcový graf, spojnicový graf nebo jejich kombinace (tzv. kombinovaný graf), výšečový graf, pruhový graf a další. Speciálními typy jsou různé teploměry, tachometry a spousta dalších.

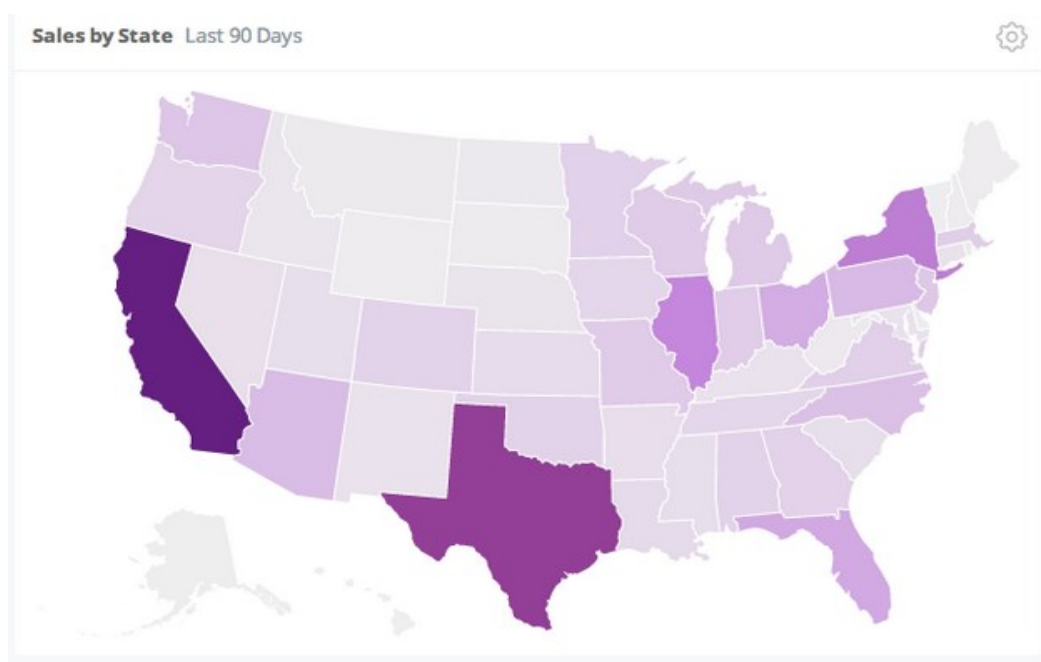


Obrázek 2-15 Ukázka tachometru
Zdroj: www.conceptdraw.com

Tabulky přehledně zobrazují data (neplatí při větším množství sloupců či řádků) a pomocí nich rychle zjistíme potřebné sumy. Někdy se používají společně s grafem, kdy tabulka zobrazí data zobrazená v grafu.

Ikony slouží především na upozornění uživatele, jestli je vše v pořádku nebo ne. Nejčastěji se využívá ikon v podobě semaforů, kdy zelená značí, že je vše v pořádku, oranžová či žlutá upozorňuje na možný problém a červená upozorňuje na určitý problém. Dalším častým způsobem jsou šipky, které ukazují výkonnost. Šipka nahoru značí nadplán, šipka rovně podle plánu a šipka dolů pod plán.

Mapy představují specializovanou vizualizaci. Umožňují přiřazení dat k určité části území (stát, kontinent, atd.). Na Obrázku 2-16 lze například vidět objem prodeje v jednotlivých státech USA.



Obrázek 2-16 Ukázka mapy
Zdroj: www.looker.com

2.9 Segment trhu Business Intelligence

V současné době se trh business intelligence velice rychle rozvíjí. Pro lepší orientaci lze rozdělit dodavatele BI do tří základních skupin (Pour, Maryška a Novotný, 2012, str. 31):

- dodavatelé komplexních řešení (s největším podílem na trhu) nabízejí celou škálu produktů včetně databázových systémů pro realizaci datových skladů a tržišť a zahrnují i všechny další nástroje potřebné pro BI řešení, jako je ETL, nástroje

pro vývoj OLAP databází atd. V této skupině figurují čtyři hlavní dodavatelé, včetně případných akvizic:

- Oracle (včetně společnosti Hyperion) s databázovým systémem Oracle;
- Microsoft s databázovým systémem MS SQL Server;
- IBM (včetně Cognos) s databázovým systémem DB2;
- SAP (včetně Business Objects);
- dodavatelé speciálních řešení a nástrojů;
- menší dodavatelé s vysoce specializovanými produkty;
- komunity a s nimi spojené společnosti poskytující open-source BI produkty.

Mezi další významné dodavatele patří například společnost HP, v oblasti datové kvality IBM, SAS, v oblasti správy metadat IBM a další.

Velmi významným trendem v rozvoji BI je integrace nástrojů a technologií BI do jiných typů aplikací, zejména ERP, kde příkladem řešení je Microsoft Dynamics AX (Axapta), které kromě transakční funkcionality poskytuje i analytické funkce na bázi využití integrovaných OLAP databází a nad nimi postavených analytických aplikací (Pour, a další, 2012).

3 Analýza současného stavu dat ve vybrané firmě

V této kapitole je představena společnost, která mi poskytla data a pro kterou bude vytvořen dashboard. Bude také popsáno, jak pracuje informační systém Evolio a jakým způsobem reportují data ze systému.

3.1 Představení společnosti

Společnost, pro kterou je navrhován dashboard se jmenuje AVE Soft s.r.o., což je česká softwarová společnost, založená v roce 1997, která se specializuje na tvorbu informačních systémů v oblasti advokacie, správy pohledávek a exekucí.

Hlavním informační systém, který společnost vyvíjí, se nazývá Evolio. Tento informační systém má několik oborových verzí. Jsou jimi Datové schránky, Advokátní kancelář, Správa pohledávek, CRM, Exekutorský úřad, Autosalón a Profi Finance. Společnost AVE Soft s.r.o. se také zabývá vývojem zakázkových mobilních aplikací.

Evolio je určeno pro všechny, kteří chtějí mít informace pod kontrolou. Společnost pracuje s Microsoft nástroji a o její vysoké úrovni a inovativnosti řešení svědčí řada úspěchů a ocenění, např. Microsoft Award Finalist 2012 nebo „Zvláštní ocenění Microsoft .NET Technology Award Finalist 2007“. Modul pro datové schránky je držitelem titulu „IT Produkt roku 2011“ časopisu ComputerWord.

3.2 Současný stav ve firmě

Informační systém pro svou práci využívá Evolio server, který zpracovává automatické úlohy. Tyto doplňky (addiny) s úlohami si může každý zákazník navolit sám. Evolio server pak na pozadí informačního systému, v určitý čas a den, spustí danou úlohu. Data, která jsou výstupem jednotlivých addinů, se posléze agregují do centrální databáze firmy. Agregace probíhá pomocí procedur, které lze vidět na Obrázku 3-1. Společnost pak na základě výsledků může informační systém vylepšovat a opravovat případně vzniklé chyby.

	Name	Description	TableName	Interval
1	AddinReport1	Chyby a varování (logy)	Profylaxe_201_AddinReport1	DAY
2	StartEndJobReport	Intervaly spuštěných jobů	Profylaxe_530_StartEndJobReport	DAY
3	DatabseReport	Profylaxe_205_DatabeseReport	Profylaxe_205_DatabaseReport	DAY
4	AddinLustrationbyTom	Statistiky lustrací	Profylaxe_530_LustrationReport	DAY
5	StartEndStepReport	Intervaly spuštěných stepů	Profylaxe_530_StartEndStepReport	DAY
6	DataboxReport	Statistika datové schránky	Profylaxe_220_DataboxReport	DAY
7	EvolioLoginsReport	Přihlášení uživatelů do evolia	Profylaxe_210_EvolioLoginsReport	MONTH

Obrázek 3-1 Procedury a jejich základní popis
Zdroj: Vlastní

Jak lze na Obrázku 3-1 vidět, agregační procedury se spouštějí většinou jednou denně. Atribut name představuje jméno procedury, description její popis, interval jak často se spouští a atribut TableName udává, do jaké tabulky se ukládají výsledky dané procedury.

V současné době však firma zobrazuje na svém interním webu pouze chyby a varování za určité období, které se dá libovolně navolit.

Nedostatkem tohoto řešení je, že tyto chyby se nedají filtrovat podle verzí Evolio serveru. Jelikož se jednotlivé verze serveru implementují zákazníkům postupně, tím že by se uváděly chyby podle verzí serveru, by bylo možné například postupné nasazení dané verze pozastavit do doby, než se objevená chyba opraví.

Tyto nedostatky se odstraní po nasazení navrhovaného dashboardu. Firma sice využívá určité grafy, ale ty jsou navrženy pomocí C#. Řešení této diplomové práce ale bude spočívat v návrhu dashboardu pomocí vybraného softwarového nástroje od firmy Microsoft, protože firma má už s Microsoft nástroji bohaté zkušenosti. Na výběr bude doplněk pro Microsoft Excel Power Pivot a Power View. Třetí možností pak bude návrh reportů pomocí návrhového prostředí v SQL Server Reporting Services (SSRS), jenž je součástí balíku Microsoft SQL Server Business Intelligence Development Studio.

4 Možnosti a návrh řešení vybraným SW nástrojem

V rámci této kapitoly bude nejprve řešena příprava zdrojových dat a testování několika nástrojů. Následně bude vybrán jeden nástroj, pomocí kterého bude dashboard vytvořen a nakonec bude otestována funkčnost navrženého dashboardu.

4.1 Příprava zdroje dat pro softwarové nástroje

Všechny tři softwarové nástroje mají stejný zdroj dat. Tedy firemní databázi, ve které jsou agregovaná data ze serverů jednotlivých zákazníků. Všechna data z Evolio serveru jsou uložena v databázi Evolio_EX_Diagnostics, která je aktualizována i několikrát denně. Firma AVE Soft s.r.o. poskytla kompletní databázi za léta 2013 a 2014 pro účely testování a vývoje dashboardu. Databáze je pro účely testování uložena na PC a po schválení od firmy se pak výsledný dashboard připojí přímo na produkční server, ze kterého jsou původní data. Veškerá příprava dat probíhala v prostředí Microsoft SQL Server a to z důvodu, že společnost AVE Soft s.r.o. taktéž tento software využívá, tím pádem se ulehčí pozdější implementace do ostrého provozu.

Pro účely testovací fáze této diplomové práce stačí pracovat z celé databáze pouze s pár tabulkami. Základem bude tabulka faktů Profylaxe_201_AddinReport1 obsahující data týkající se počtu chyb, kritických chyb, varování atd. určitých stepů jednotlivých verzí Evolio serveru (Obrázek 4-1). Na tuto tabulku je navázána, pomocí cizích klíčů, tabulka ProfylaxeReport a na ní pak tabulka ProfylaxeServer.

	IDReportAddin	IDReport	Job	Step	RunCount	FinishCount	ExceptionCount	LogWarningCount	LogErrorCount	LogCriticalErrorCount	Datum
1	3980	15093	Export do CEE - Fáze 2 - t...	CEEExport_Phase2	1	0	1	1	6	1	2013-09-13
2	6457	15094	CEE export 1 - vytvoření ...	CEEExport_Phase2	1	0	1	114	0	1	2013-09-06
3	13992	15298	Export do CEE - Fáze 2 - t...	CEEExport_Phase2	1	0	1	2	6	1	2013-09-27
4	17245	15414	CUZK 2 - lustrace on-line	CUZK_Phase2	2	0	2	0	0	2	2013-10-07
5	18023	15445	CEE export 1 - vytvoření ...	CEEExport_Phase2	1	0	1	96	1	1	2013-10-07
6	18207	15447	CUZK 2 - lustrace on-line	CUZK_Phase2	4	1	3	0	0	3	2013-10-08
7	23553	15697	Export do CEE - Fáze 2 - t...	CEEExport_Phase2	1	0	1	0	6	1	2013-10-25

Obrázek 4-1 Ukázka dat z tabulky faktů Profylaxe_201_AddinReport1
Zdroj: Vlastní

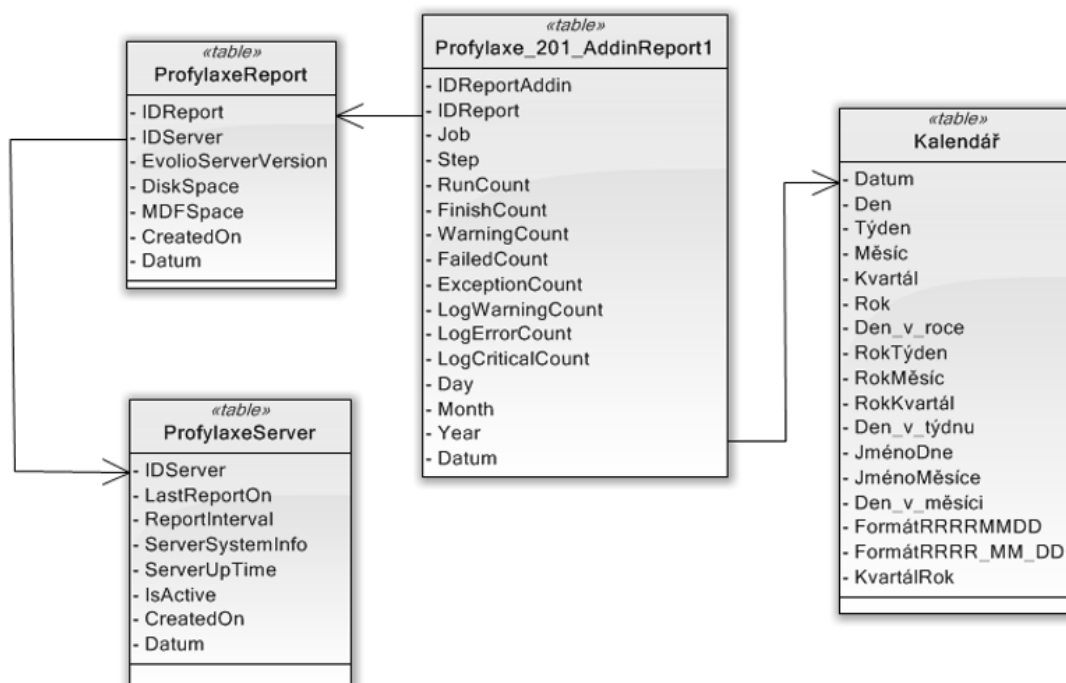
Jelikož databáze neobsahovala dimenzi času, bylo potřeba ji vytvořit (Obrázek 4-2). Po vytvoření byl kalendář pomocí cizího klíče propojen s faktovou tabulkou Profylaxe_201_AddinReport1, bylo ovšem nutné přidat další sloupec (atribut) s názvem Datum obsahující datum s datovým typem date. Kalendář obsahuje data za roky 2000 až 2050.

	Datum	Den	Týden	Měsíc	Kvartál	Rok	Den_v_roce	RokTýden	RokMěsíc	RokKvartál	Den_v_týdnu	JménoDne	JménoMěsíce	Den_v_měsíci	FormatRRRRMMDD	FormatRRRR_MM_DD
1	2000-01-01	1	1	1	1	2000	1	200001	200001	200001	6	Saturday	January	31	20000101	2000-01-01
2	2000-01-02	2	1	1	1	2000	2	200001	200001	200001	7	Sunday	January	31	20000102	2000-01-02
3	2000-01-03	3	2	1	1	2000	3	200002	200001	200001	1	Monday	January	31	20000103	2000-01-03
4	2000-01-04	4	2	1	1	2000	4	200002	200001	200001	2	Tuesday	January	31	20000104	2000-01-04
5	2000-01-05	5	2	1	1	2000	5	200002	200001	200001	3	Wednesday	January	31	20000105	2000-01-05
6	2000-01-06	6	2	1	1	2000	6	200002	200001	200001	4	Thursday	January	31	20000106	2000-01-06
7	2000-01-07	7	2	1	1	2000	7	200002	200001	200001	5	Friday	January	31	20000107	2000-01-07

Obrázek 4-2 Ukázka dat v tabulce Kalendář
Zdroj: Vlastní

Kalendář byl vytvořen hlavně z důvodu filtrování podle času v dashboardu, což by bez kalendáře nebylo možné.

Pro přehlednost je na Obrázku 4-3 uvedeno schéma tabulek připravené pro testování. Jak už bylo řečeno, toto schéma bude pro všechny tři softwarové nástroje totožné.



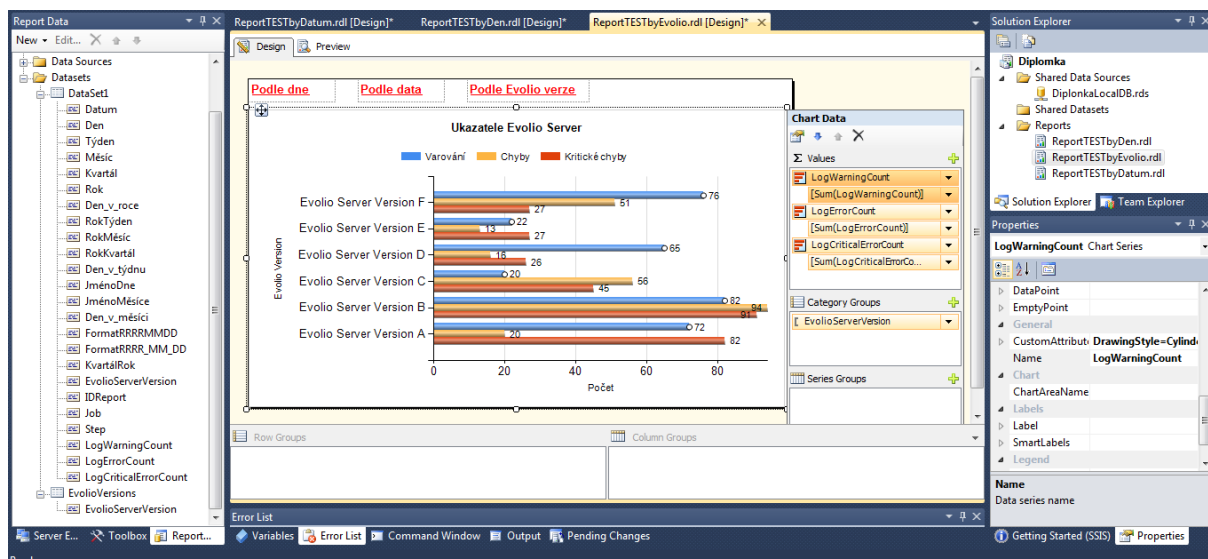
Obrázek 4-3 Schéma tabulek pro testování nástrojů
Zdroj: Vlastní

4.2 SQL Server Reporting Services (SSRS)

Prvním zvoleným nástrojem pro vytvoření dashboardu je SQL Server Reporting Services (SSRS) z balíčku SQL Server Business Intelligence Development Studio.

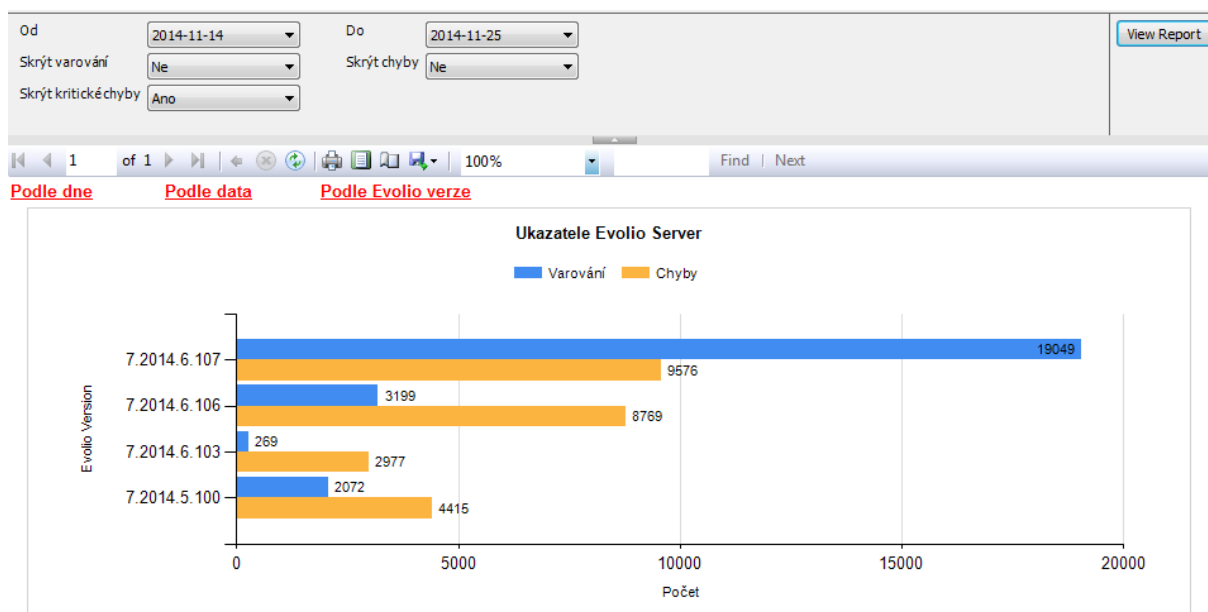
Základním krokem vytvoření dashboardu je tvorba jednotlivých reportů, které čtou data z datového skladu obsahující tabulky faktů i jednotlivé dimenzní tabulky.

Základem budoucího dashboardu je projekt, který bude obsahovat všechny reporty obsažené na případném dashboardu. Dalším krokem je vytvoření datového zdroje, jinak řečeno připojení do databáze, ze kterého se budou čerpat data. Těchto zdrojů může být v rámci jednoho projektu hned několik, tudíž reporty mohou být v jednom projektu vytvořeny z různých databází. Zdrojem může být jak relační databáze tak i multidimenzionální. Ukázkou vytvořeného reportu, včetně rozhraní, lze vidět na Obrázku 4-4.



Obrázek 4-4 Uživatelské prostředí SSRS
Zdroj: Vlastní

Reporty vytvořené pomocí tohoto nástroje se pak následně umístí (tzv. deploy) na Microsoft Office Sharepoint Server. Na tomto serveru se vytvoří webová stránka obsahující menu, pomocí kterého si lze zobrazovat příslušné reporty. Na každé stránce se pak zvolí report, který má být zobrazen. V tomto případě tedy dashboard představuje jakousi webovou stránku se záložkami obsahující reporty.



Obrázek 4-5 Ukázka výsledného reportu
Zdroj: Vlastní

Výhodou tohoto řešení je, že výsledný dashboard je dostupný na podnikovém portálu a uživatelé, pokud mají přidělená práva, si ho tedy mohou kdykoliv zobrazit. Ovšem nevýhodou je, že uživatel může pouze filtrovat a není zde žádná možnost přizpůsobení uživatelským požadavkům či změny ukazatelů atd. Report je až na filtrování statický. Dalším neduhem

je rychlost načítání reportu, které se s větším počtem tabulek a dat zpomaluje. Problém tkví taktéž v tom, že při každé změně filtrování se report znovu celý načítá.

Co se týče pořizovacích nákladů, pohybují se v desítkách tisíců za licenci jak na zmíněný celý balík, tak i na Microsoft Sharepoint plus náklady za servis a server, na kterém by Sharepoint běžel.

4.3 Power Pivot

Druhým testovaným nástrojem je Power Pivot, což je doplněk Microsoft Excelu. Funguje jak ve verzi 2010, tak i ve 2013. Tento doplněk je zdarma a podle distribuční verze lze tento doplněk buď aktivovat přímo v nastavení Excelu, nebo stáhnout z oficiálních stránek Microsoftu a poté nainstalovat. Power Pivot v podstatě posouvá BI do prostředí Excelu a je z pohledu licenční politiky daleko levnější než jiné nástroje, které jsou většinou těžkopádnější. Cena licence se pohybuje v jednotkách tisíc za Microsoft Excel a samotný doplněk je zadarmo.

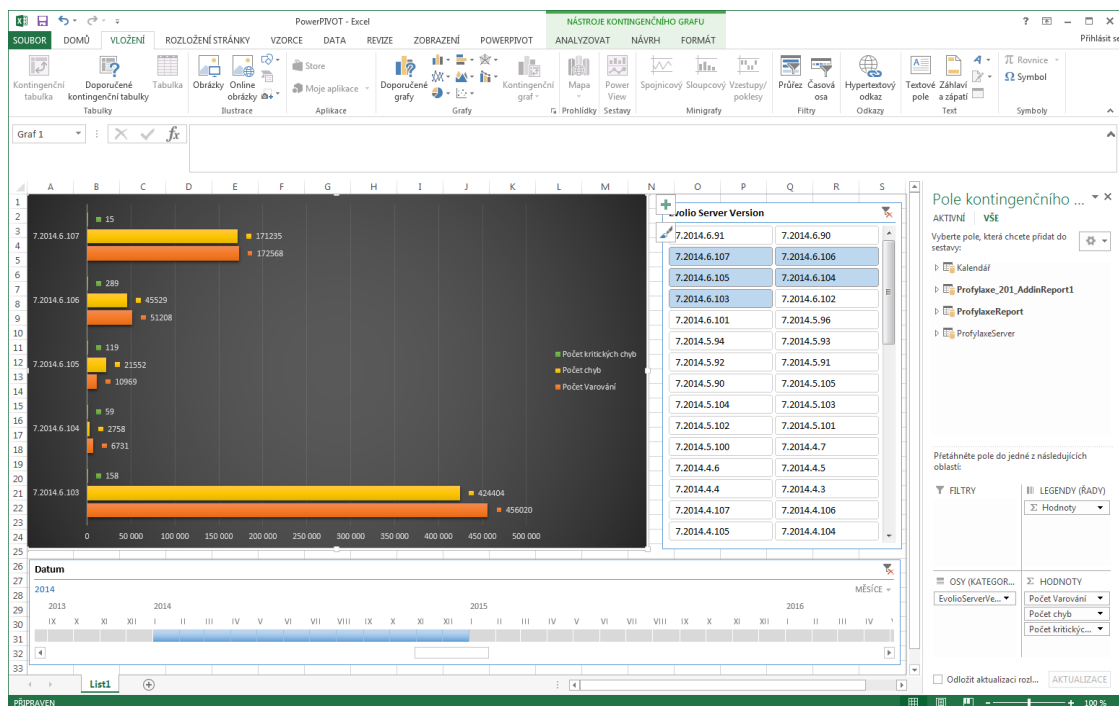
Oproti klasickému Excelu nejsou data uložena přímo v sešitu, ale v samostatné databázi doplňku Power Pivot, který data ukládá do relační databáze. Power Pivot používá pro dotazování do dat jazyk DAX. Umožňuje vytváření hierarchií, což se jinak v relační databázi nedá. Nevýhodou Power Pivotu týkající se pouze 32-bitové verze je omezení přibližně jednoho milionu řádků na sešit, což v případě obrovského množství dat v zdrojových tabulkách způsobuje problémy.



Obrázek 4-6 Hierarchie Rok->Den
Zdroj: Vlastní

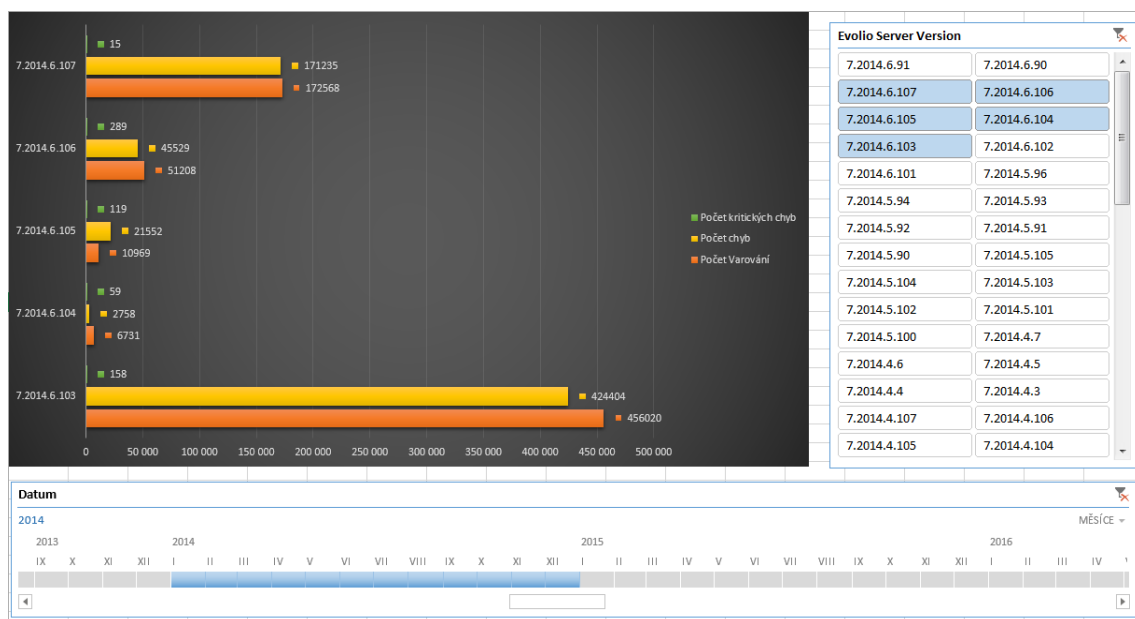
Před vytvářením reportů bylo nutné vytvořit hierarchii Rok->Den umožňující tzv. „drill-down“ a „drill-up“ po těchto úrovních Rok >> Kvartál >> Měsíc >> Týden >> Den.

Ukázka reportu s využitím hierarchie lze vidět na Obrázku 4-7. Hierarchie byla využita pro zobrazení časové osy umožňující zobrazovat datum detailněji.



Obrázek 4-7 Uživatelské prostředí doplňku Power Pivot
Zdroj: Vlastní

Na Obrázku 4-7 jde taktéž vidět kompletní uživatelské prostředí, které uvidí uživatel. Zde si podle potřeby může koncový uživatel přidávat jednotlivé atributy na osy x a y, filtry a hodnoty. Pokud je toto uživateli umožněno, vyvstává otázka, zda stačí vytvořit jeden report, který si uživatel přizpůsobí sám nebo se vytvoří více reportů, které budou rozděleny po záložkách v Excelu a budou tak představovat dashboard, kde si uživatelé budou filtrovat podle potřeby. Druhá možnost je samozřejmě praktičtější a pro rychle zkontrolování údajů určitě efektivnější.



Obrázek 4-8 Ukázka reportu s využitím hierarchie navrženého pomocí Power Pivot
Zdroj: Vlastní

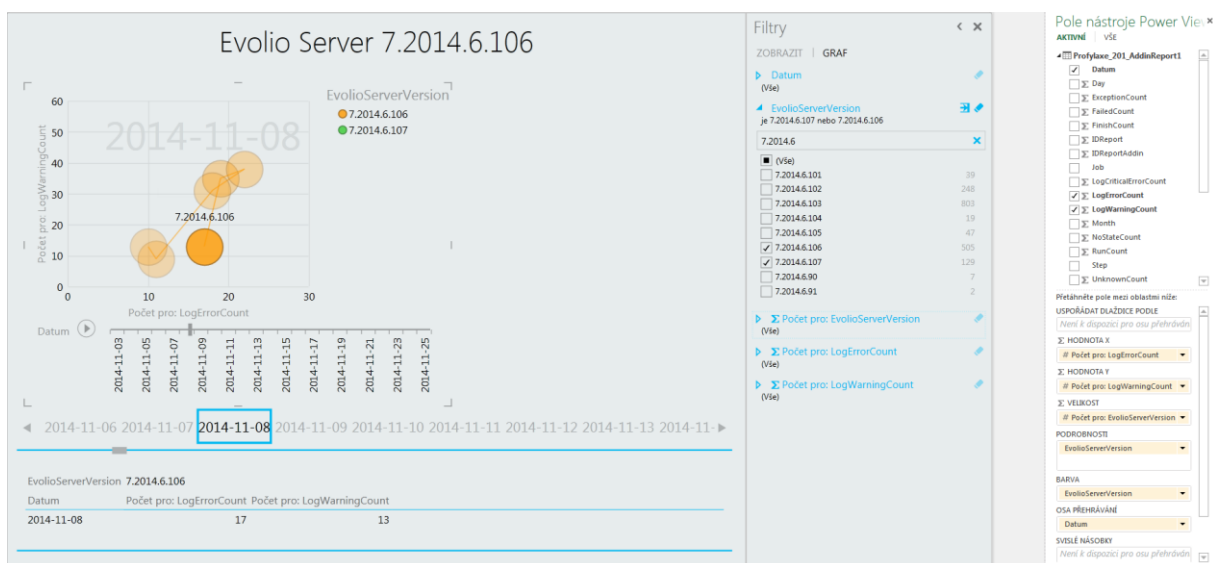
Na příkladu uvedeném na Obrázku 4-8 lze pozorovat počet chyb a počet varování u verzí Evolio Serveru 7.2014.6.104 až 7.2014.6.107 za celý rok 2014. Tyto ukazatele lze v pravém panelu velice jednoduše měnit přetažením pole do určité oblasti. Prostředí je velice intuitivní a flexibilní. Power Pivot umožňuje vytvářet hierarchie, kalkulované sloupce atd.

Pomocí jazyka DAX lze získávat hodnoty, které je možné využít v sestavách (reportech). Výrazy jazyka DAX jsou v mnohém podobné vzorcům Excelu. Tento jazyk je ovšem už pro zkušenější uživatele a pro potřeby této diplomové práce není potřeba jeho funkčnosti využít.

4.4 Power View

Posledním testovaným nástrojem je doplněk Excelu Power View. Podle distribuční verze je možné tento doplněk, stejně jako u Power Pivotu, buď aktivovat v nastavení Excelu, nebo stáhnout z oficiálních stránek Microsoftu a posléze nainstalovat.

Power View je nástroj pro interaktivní zkoumání, vizualizaci a prezentaci dat. Zdrojem dat mohou pocházet z datového modelu Power Pivotu nebo přímo z relační databáze.



Obrázek 4-9 Ukázka Power View
Zdroj: Vlastní

Na Obrázku 4-9, zachycující příklad z doplňku Power View, lze vidět uživatelské prostředí. Graf zobrazuje historii chybovosti a varování z Evolio verze 7.2014.6.106. Pouze tento typ grafu obsahuje spustitelnou časovou osu, jež velice přehledně ukazuje vývoj ukazatelů Evolio serveru 7.2014.6.106. V pravém panelu si lze podle potřeby graf přizpůsobit. Výhodou je již dříve zmíněná časová osa. Pokud zdroj dat obsahuje státy, lze v tomto doplňku využít mapy, na kterých se stejně jako na grafu, zobrazí bubliny znázorňující zvolené ukazatele.

Ovšem prostředí tohoto doplňku už tak intuitivní není. Pás karet je nelogicky uspořádán a uživatel neví, co k čemu patří. Nelze měnit názvy atributů a doplněk občas přestane pracovat a je nutný restart. Power View je relativně nový doplněk, je tedy možné, že se časem zmíněné neduhy aktualizacemi opraví a přidá nové funkčnosti.

4.5 Zvolení SW nástroje

V následující tabulce jsou uvedeny výhody a nevýhody jednotlivých nástrojů. Tyto výhody a nevýhody se vztahují pouze na to, když je zdrojem dat relační databáze. Pokud by byla zdrojem dat multidimenzionální databáze, hierarchie a kalkulované sloupce by už byly v databázi předpřipravené, tudíž by nemělo smysl řešit, zda to nástroje zvládají vytvořit sami či ne.

Nástroj	Výhody	Nevýhody
SSRS	<ul style="list-style-type: none"> - Portálové řešení - Dostupnost pro uživatele - Použití více zdrojů dat najednou 	<ul style="list-style-type: none"> - Tvorba reportů pouze pokročilým uživatelem - Nutnost mít zároveň Microsoft Sharepoint - Pomalé načítání reportů - Znovu načítání celého reportu při každé změně při filtrování - Cena (je nutné mít licenci jak na SQL Server Business Intelligence Development Studio tak i na Microsoft Sharepoint)
Power Pivot	<ul style="list-style-type: none"> - Uživatelské prostředí - Flexibilita návrhu reportů - Uživatelsky přívětivý nástroj - Rychlost - Tvorba hierarchií - Tvorba kalkulovaných sloupců - Dostupné v Excelu 2010 i 2013 - Použití pouze jednoho zdroje dat - Zdarma 	<ul style="list-style-type: none"> - Limit cca 1 milion řádků na sešit - Nelze použít speciální grafy (teploměr, tachometr,...)
Power View	<ul style="list-style-type: none"> - Grafické prostředí - Spustitelná časová osa - Mapy - Zdarma - Power Pivot jako zdroj 	<ul style="list-style-type: none"> - Spustitelná časová osa pouze u jednoho typu grafu (Bubble graf) - Dostupnost pouze v Excelu 2013 - Nedodělanost doplňku - Uživatelsky nepřívětivé prostředí a ovládání - Buggy - Nemožnost tvorby hierarchií a kalkulovaných sloupců - Nemožnost měnit název atributu

Tabulka 4-1 Porovnání nástrojů
Zdroj: Vlastní

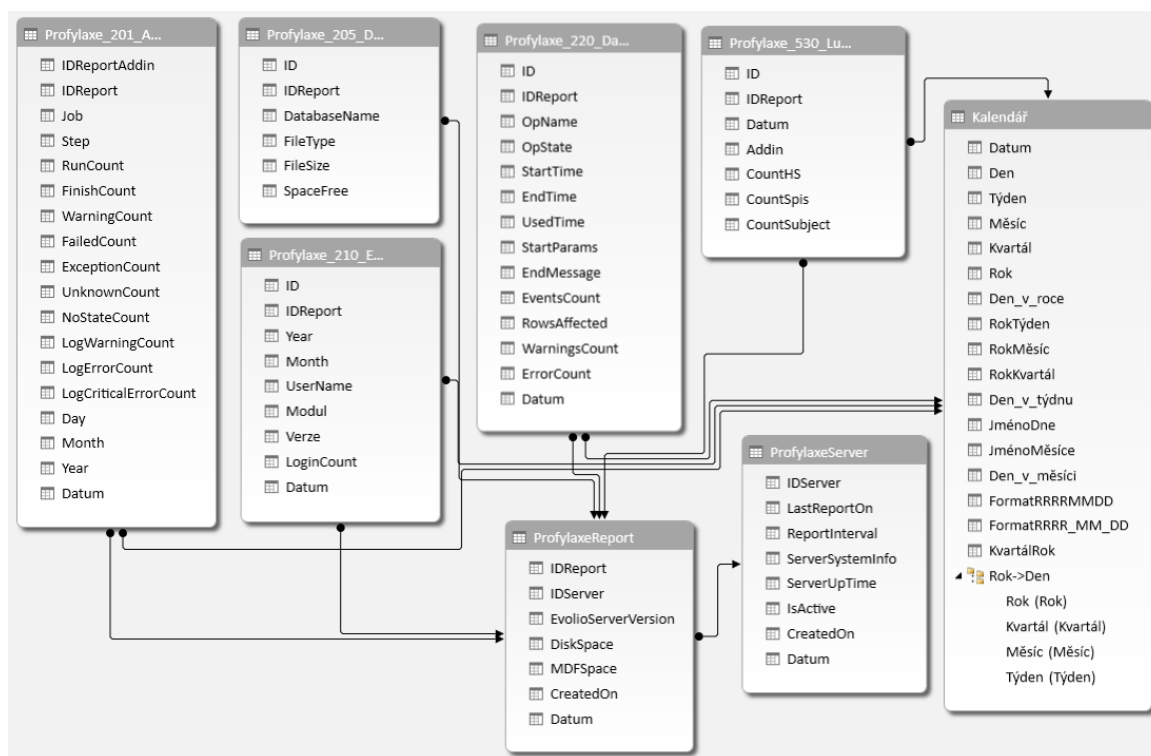
Už podle tabulky lze vyvodit, že nejlepší z vyzkoušených nástrojů je Power Pivot obsahující v drtivé většině samá pozitiva. Obecné výhody a nevýhody už byly zmíněny. Co se týče vhodnosti nasazení do společnosti AVE Soft s.r.o., tak je taktéž doporučeno použít nástroj Power Pivot. Firma sice používá Microsoft Sharepoint, ale vzhledem k tomu, že výsledný dashboard bude sloužit pouze pár zaměstnancům firmy, kteří si budou navíc moci jednotlivé reporty přizpůsobit svým potřebám, jeví se použití SSRS jako nevhodné řešení. To samé by se dalo říci i o nástroji Power View, jenž je pro uživatele těžkopádný na ovládání a při testování vykazoval značnou nestabilitu a omezenou funkčností.

4.6 Návrh a testování dashboardu

V předchozí kapitole byl pro vytvoření dashboardu zvolen nástroj Power Pivot. Jak už bylo zmíněno dříve, porovnání probíhalo na testovací databázi. Při vytváření dashboardu už ovšem bude čerpáno přímo z produkční databáze, aby se nemusela v budoucnu řešit konektivita. Úpravy provedené v testovací databázi se pomocí skriptu aplikují i do databáze produkční.

4.6.1 Datový model

Jelikož dashboard bude čerpat i z dalších faktových tabulek, bylo nutné kalendář na tyto nové faktové tabulky pomocí cizích klíčů a v některých případech i nově vytvořených atributů připojit. Finální datový model, ze kterého vychází dashboard, je zobrazen na Obrázku 4-10.





Obrázek 4-10 Datový model zdrojového datamartu

Datový model má podobu sněhové vločky a obsahuje pět faktových tabulek. Profylaxe_201_AddinReport1 udává chyby a varování Evolio serveru. Profylaxe_205_DatabaseReport obsahuje informace o jednotlivých databázích na serverech. Profylaxe_210_EvolioLoginsReport eviduje přihlášení uživatelů do jednotlivých modulů informačního systému Evolio. Profylaxe_220_DataboxReport zobrazuje statistiky datových schránek a Profylaxe_530_LustrationReport zase statistiky lustrací.

4.6.2 Testovací verze dashboardu

Obrázky níže obsahují jednu z prvních verzí dashboardu, která byla poskytnuta společnosti Ave Soft s.r.o., jejíž poznatky pak byly zakomponovány do finální verze dashboardu.

V prvním sešitě se nachází úvodní stránka (Obrázek 4-11), kde je uveden stručný popis, zdroj dat, ze kterých dashboard čerpá a popis jednotlivých záložek představující vždy jednu faktovou tabulku.

EVOLIO dashboard

Popis:

Tento dashboard slouží pro analýzu ukazatelů EVOLIO serveru. Mezi tyto ukazatele patří chyby a varování v systému, volné místo na disku serveru, přihlašování uživatelů IS, statistiky lustrací a datových schránek atd. Jednotivé reporty se pak nacházejí v záložkách, které jsou popsány v další sekci. Každý sešit obsahuje jednu faktovou tabulku.

Vstupní data:

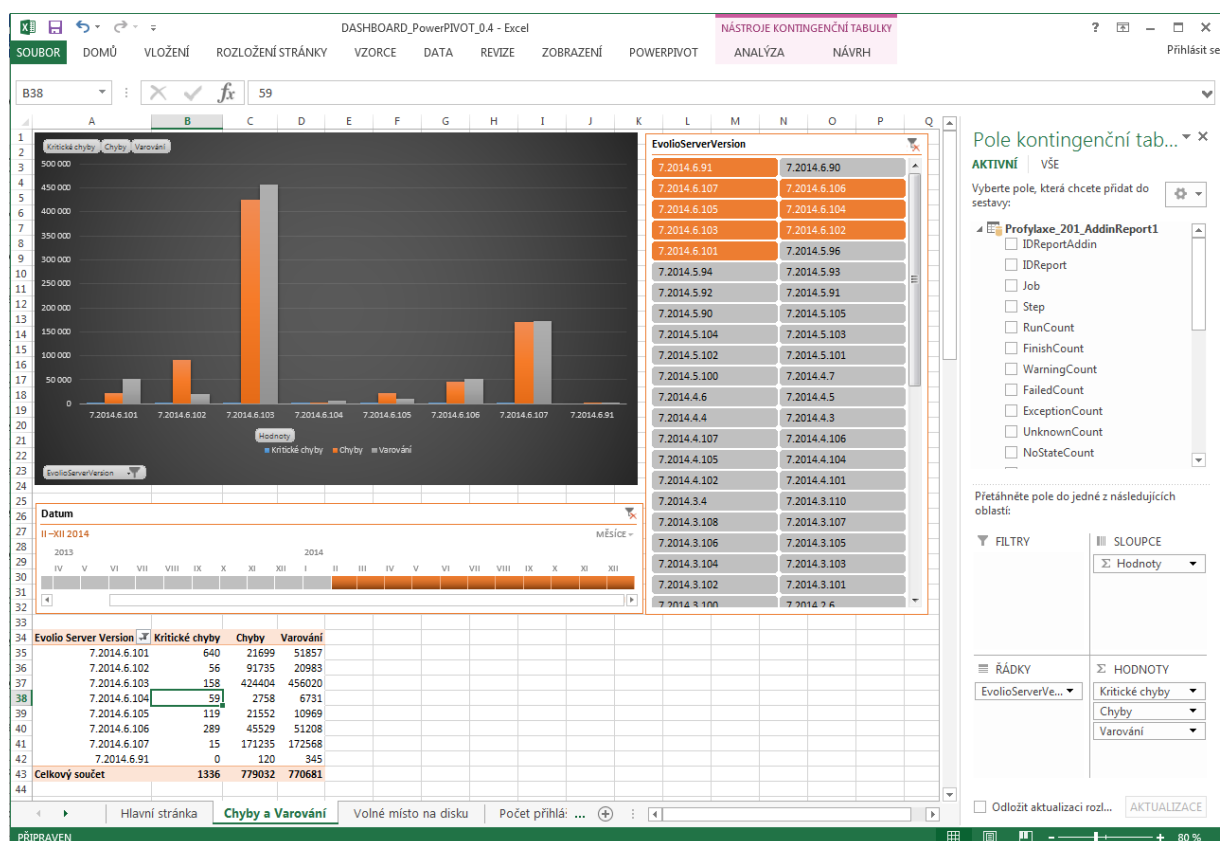
Data pocházejí z produkčního serveru firmy AVE Soft s.r.o., kde zdrojovou datbází je Evolio_EX_Diagnostic. Faktovými tabulkami použitými v tomto dashboardu jsou: Profylaxe_201_AddinReport1, Profylaxe_205_DatabaseReport, Profylaxe_210_EvolioLoginsReport, Profylaxe_220_DataboxReport a Profylaxe_530_LustrationReport.

Záložky:

Hlavní stránka:	Popis dashboardu, vstupních dat a jednotlivých záložek obsahují reporty
Chyby a Varování	Kritické chyby, chyby a varování (logy)
Volné místo na disku	Volné místo v jednotlivých databázích na serveru
Počet přihlášení do IS	Počet přihlášení do jednotlivých modulů systému
Datové schránky	Informace o datových schránkách (chyby, varování, počet událostí atd.)
Lustrace	Počet HS, spisů a předmětů v jednotlivých oddílech

Obrázek 4-11 Úvodní stránka
Zdroj: Vlastní

Na Obrázku 4-12 lze vidět ukázkou jednoho ze sešitů dashboardu. Tato ukáзка obsahuje graf, který je přímo propojen s tabulkou pod ním. Pod grafem je umístěna časová osa pro velice pohodlné filtrování podle času. Vedle grafu je zobrazen seznam verzí Evolio serveru pomocí něhož lze filtrovat údaje v grafu respektive v tabulce. Pravý panel slouží pro přidávání nových atributů či ukazatelů.



Obrázek 4-12 Ukázka vybrané záložky dashboardu
Zdroj: Vlastní

V prvních fázích vývoje je každý ze sešitů dashboardu proveden v jiném stylu, ze kterého si firma mohla vybrat jeden z nich či ponechat dashboard tak jak je.

4.6.3 Testování

Testovací verze dashboardu, stejně tak jako finální, je vlastně jeden “excelovský” soubor. Testovací verze byla poskytnuta pár lidem, kteří rozhodnou o tom, zda jsou jednotlivé záložky dobře rozvrženy, jestli obsahují správné ukazatele a hlavně otestovali, zda data odpovídají realitě. Součástí testování je i zvolení vhodného vzhledu celého dashboardu.

Výstupem praktické části diplomové práce je dashboard, který je sice dokončený, ale bohužel z důvodu velké vytíženosti a nedostatku času zaměstnanců firmy není firmou oficiálně schválený a uvedený do ostrého provozu.

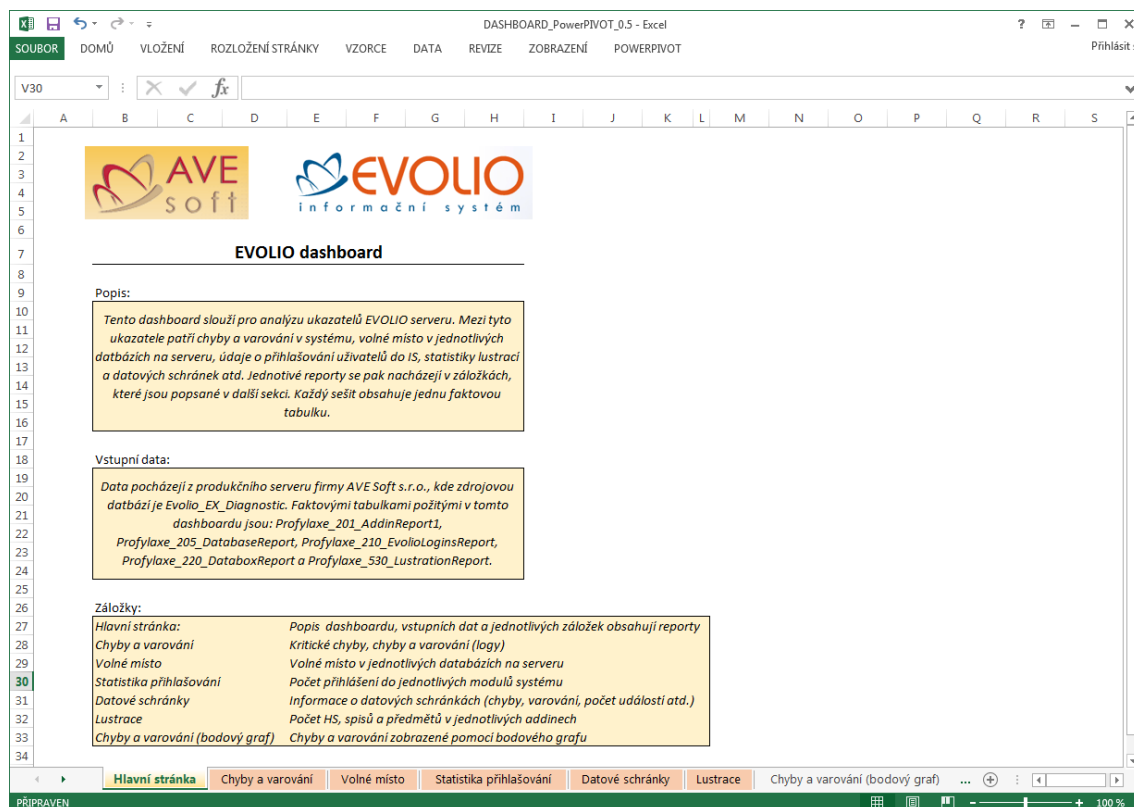
4.7 Finální verze dashboardu

V následujících odstavcích je shrnutý obsah jednotlivých záložek dashboardu. Tato verze v sobě zahrnuje veškeré dosud získané poznatky, změny a návrhy od uživatelů a autora práce, kteří dashboard testovali. První změnou, oproti testovací verzi, je sjednocení vzhledu, druhou je přidání nové záložky využívající doplněk Power View čerpající data

z Power Pivotu a poslední změnou pak viditelnost, respektive neviditelnost některých sloupců, které jsou zbytečné, aby se v uživatelském prostředí zobrazovaly.

Kromě hlavní stránky je ve všech záložkách k dispozici časová osa, pomocí které lze velice jednoduše data filtrovat.

Hlavní stránka

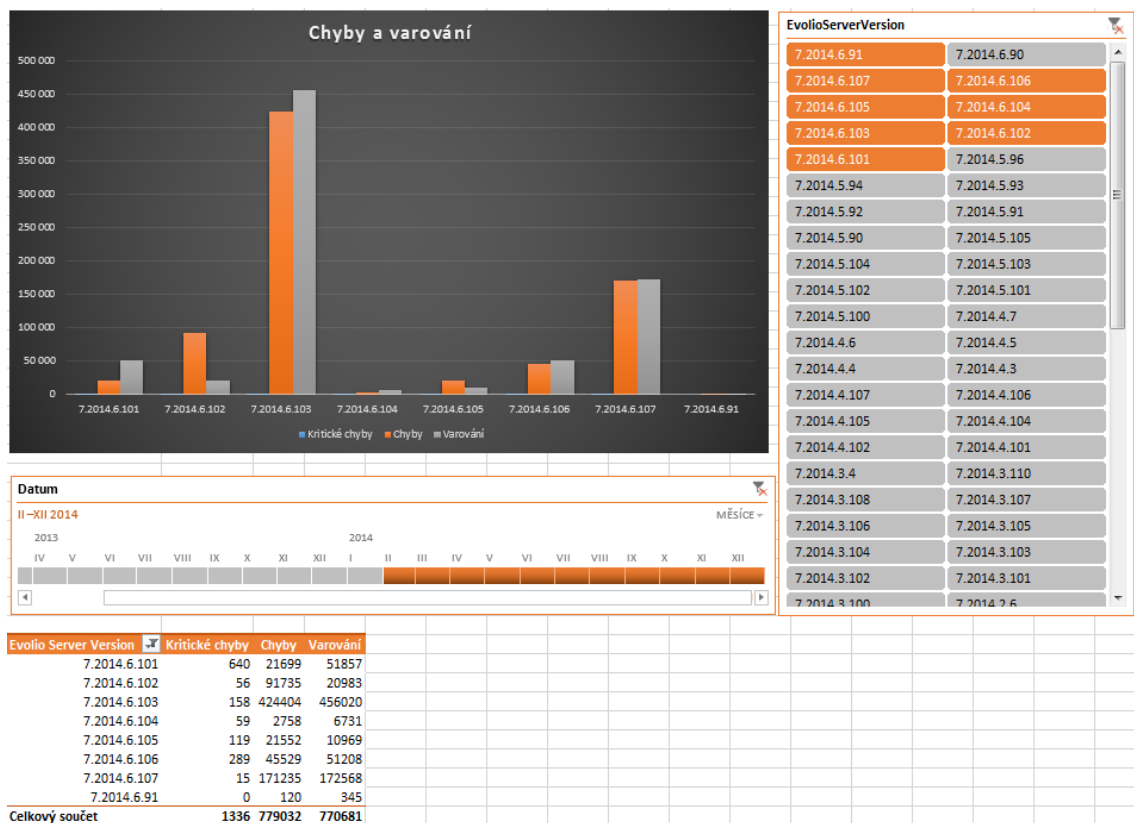


Obrázek 4-13 Hlavní stránka dashboardu včetně uživatelského prostředí
Zdroj: Vlastní

Na Obrázku 4-13 lze vidět uživatelské prostředí včetně hlavní stránky dashboardu. Na spodní liště pak lze vidět jednotlivé záložky, které jsou dále postupně popsány. Na hlavní stránce je uveden stručný popis dashboardu spolu s výčtem základních ukazatelů. Dále jsou zde uvedena vstupní data, kde je napsána zdrojová databáze a její tabulky, které celý dashboard používá. Na hlavní stránce taktéž najdeme názvy záložek a jejich obsah.

Chyby a varování

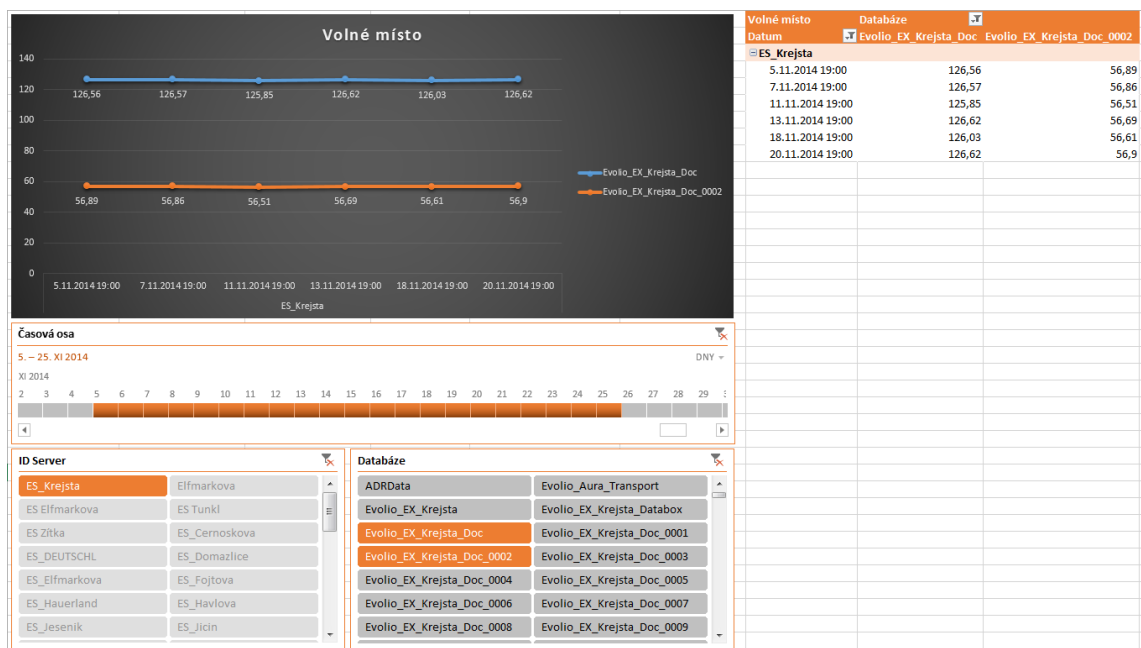
Záložka chyby a varování (Obrázek 4-14) zobrazuje základní ukazatele, které generují procedury mající na starost zejména chyby a varování vyskytující se při provádění určitých jobů v Evolio serveru. Tyto chyby a varování jsou pak rozděleny podle verzí Evolio serveru. Data v grafu a tabulce jsou totožná. Uživatel má pak ještě na výběr například ukazatel kritických chyb a podrobnosti všech ukazatelů si může rozdělit po jednotlivých jebch.



Obrázek 4-14 Chyby a varování
Zdroj: Vlastní

Volné místo

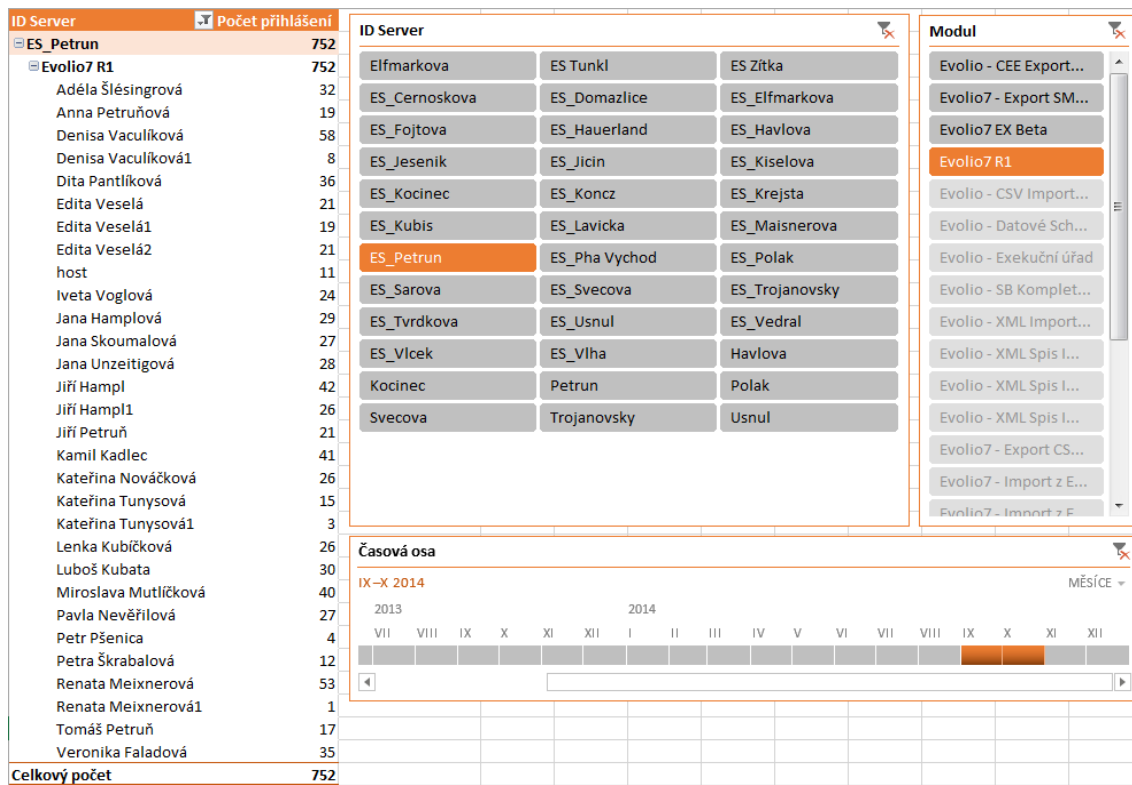
Už podle názvu, tato záložka zobrazuje, kolik zbývá volného místa v jednotlivých databázích na serveru zákazníka. Data v grafu jsou totožná s daty v tabulce. Filtrování s databázemi se mění v závislosti na vybraném serveru.



Obrázek 4-15 Volné místo
Zdroj: Vlastní

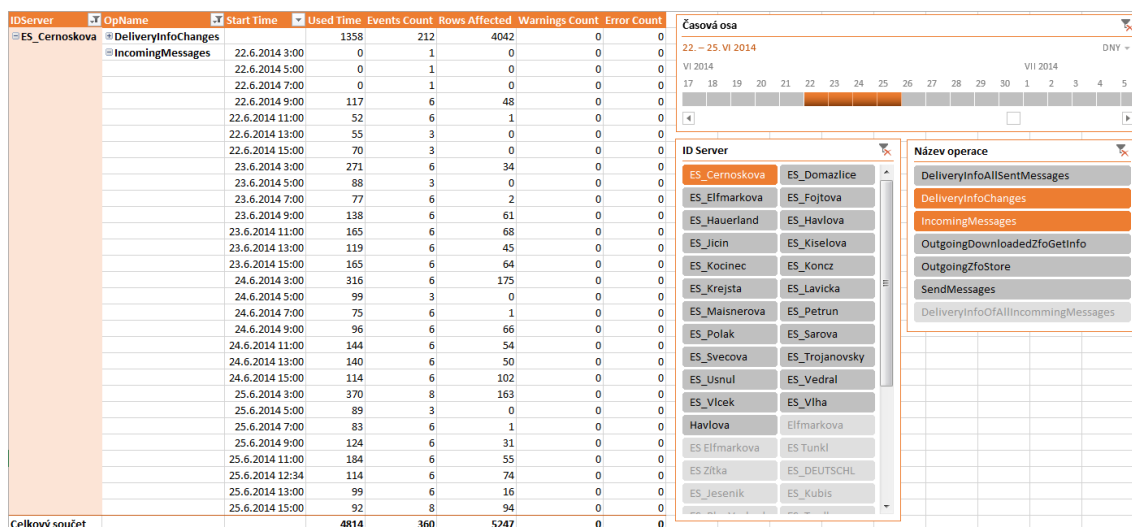
Statistika přihlašování

Tato záložka zobrazuje statistiku přihlašování do jednotlivých modulů informačního systému Evolio. Stejně jako u předchozí záložky i zde se mění moduly v závislosti na vybraném serveru. Kvůli přehlednosti byl vynechán atribut i filtr verze Evolio serveru, které si může uživatel samozřejmě přidat.



Obrázek 4-16 Statistika přihlašování
Zdroj: Vlastní

Datové schránky

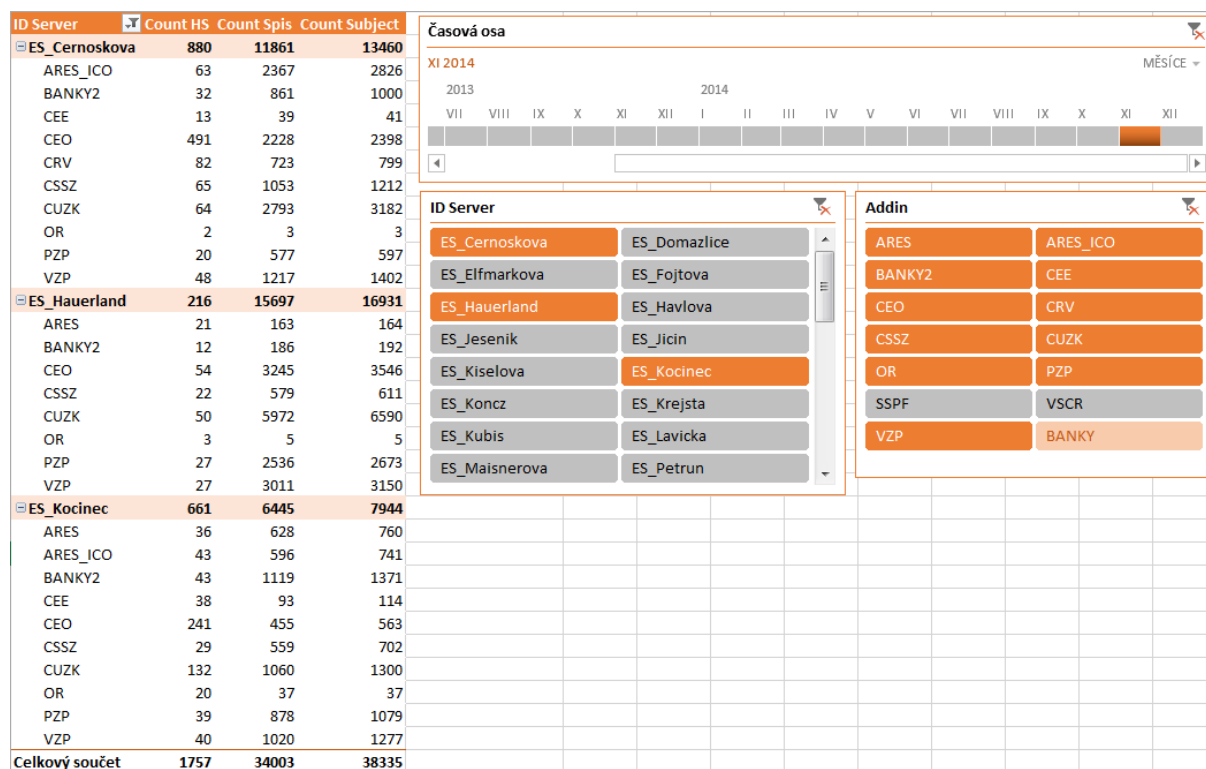


Obrázek 4-17 Datové schránky
Zdroj: Vlastní

Záložka datové schránky zobrazuje informace, jakými jsou například chyby, varování, počet událostí atd. Filtrovat lze podle typu (názvu) operace, jejíž výběr se mění v závislosti na vybraném serveru. Ukázka je vidět na Obrázku 4-17.

Lustrace

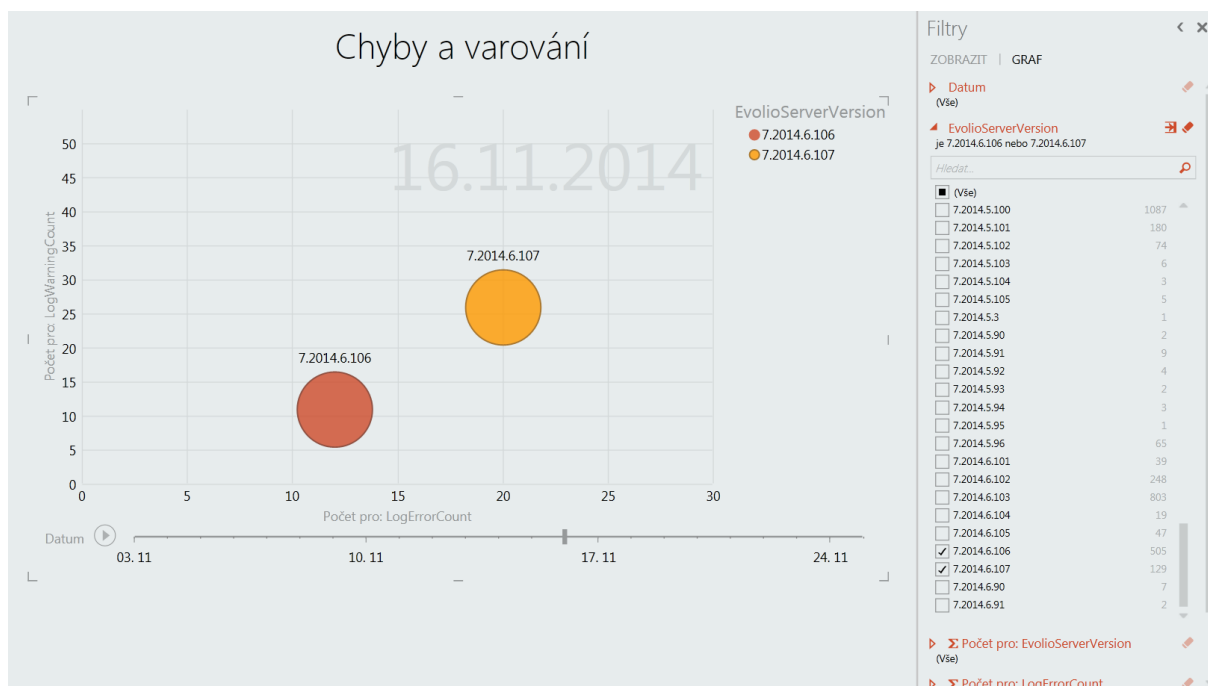
V této záložce (Obrázek 4-18) jsou uvedeny základní ukazatele týkající se lustrací. Lze filtrovat podle addinů na základě výběru serveru či serverů.



Obrázek 4-18 Lustrace
Zdroj: Vlastní

Chyby a varování (bodový graf)

Tato speciální záložka je na rozdíl od předcházejících vytvořena pomocí doplňku Power View, který nabízí spustitelnou časovou osu umožňující v tomto případě vidět v bodovém grafu vývoj chyb a varování vybraných Evolio serverů v čase.



Obrázek 4-19 Chyby a varování (bodový graf)
Zdroj: Vlastní

4.8 Shrnutí

V praktické části je zachycen celý postup, který vedl k vytvoření dashboardu. Z počátku bylo nutné analyzovat data a podle toho se rozhodnout, co bude dashboard obsahovat. Poté následovala fáze výběru vhodného nástroje pro vytvoření dashboardu. Zde bylo na výběr ze tří Microsoft nástrojů a to SSRS, Power View a Power Pivot. Po otestování všech možností byla vybrána kombinace Power Pivotu a Power View, kde právě jedna záložka je vytvořena pomocí doplňku Power View a zbývající pak v doplňku Power Pivot. Vytvořený dashboard byl poskytnut firmě k otestování. Dashboard je kompletně vytvořený, ale z důvodu nedostatku času ze strany firmy už bohužel nedošlo k nasazení dashboardu do ostrého provozu.

5 Závěr

Dashboards jsou dnes většinou nepostradatelnou pomocí každého manažera, který na základě přehledně zobrazených dat v dashboardu dokáže velice rychle učinit určitá rozhodnutí. Diplomová práce byla nejdřív zaměřena na analýzu toho, jak firma reportuje data ze svého informačního systému. Poté byla věnována pozornost výběru vhodného nástroje pro tvorbu dashboardu a jeho následné vytvoření pomocí zvoleného nástroje.

Cílem práce bylo vytvoření dashboardu pro firmu AVE Soft s.r.o., která by byla z navrženého dashboardu schopna vyčíst ukazatele informačního systému Evolio a podle toho například ovlivnit nasazování nových verzí systému jednotlivým zákazníkům.

Mezi zmíněné ukazatele se například řadí chyby a varování v systému, volné místo na disku serveru, přihlašování uživatelů do jednotlivých modulů IS, statistiky lustrací a datových schránek atd. Pro vytvoření dashboardu byly vyzkoušeny tři různé nástroje od firmy Microsoft, z nichž nakonec byla použita kombinace nástrojů Power Pivot a Power View. V době odevzdání této práce byl dashboard již kompletně vytvořen, ale ještě se nenacházel v ostrém provozu. Nicméně cíl práce, tedy vytvoření dashboardu, byl splněn.

Problematika Business Intelligence byla nastíněna v teoretické části této diplomové práce. Základní teorie a technologie včetně celé řady pojmů s BI souvisejících vychází z prostudované odborné literatury.

V analytické části pak byla nejprve představena společnost a následně popsán proces, jakým způsobem informační systém Evolio pracuje a jak z něj firma reportuje data.

Jednotlivé kroky, které vedly k návrhu kompletního dashboardu pro firmu AVE Soft s.r.o., byly popsány v praktické části práce.

Co se týče možností dalšího vývoje dashboardu, tak by po nasazení do ostrého provozu mohla být vytvořena aplikace pro mobilní platformy zahrnující autorem vytvořený dashboard. Chytré telefony a tablety jsou dnes již nedílnou součástí každého z nás a dashboard by tak byl pro zaměstnance, zejména pro manažery firmy, kdykoliv po ruce a mohli by tak činit rozhodnutí i když by se zrovna nenacházeli u svého počítače.

Seznam použité literatury

ECKERSON, Wayne W. *Performance dashboards: measuring, monitoring, and managing your business*. 2nd. New York: Wiley, 2006. ISBN 978-047-0589-830.

FEW, Stephen. *Information Dashboard Design: The Effective Visual Communication of Data*. Sebastopol: O'Reilly, 2006. ISBN 978-0-596-10016-2.

INMON, William H. *Building the data warehouse*. 4th. Indianapolis: Wiley, 2005. ISBN 978-0764599446.

LACKO, Ľuboslav. *Business Intelligence v SQL Serveru 2008: reportovací, analytické a další datové služby*. Brno: Computer Press, 2009. ISBN 978-80-251-2887-9.

LARSON, Brian. *Delivering business intelligence with Microsoft SQL server 2012*. 3rd. New York: McGraw-Hill, 2002. ISBN 00-717-5938-7.

LAURENČÍK, Marek a Michal BUREŠ. *Excel 2013: práce s databázemi a kontingenčními tabulkami*. Praha: Grada, 2014. ISBN 978-80-247-5003-3.

NOVOTNÝ, Ota. *Business intelligence: jak využít bohatství ve vašich datech*. Praha: Grada, 2005. ISBN 80-247-1094-3.

POUR, Jan; MARYŠKA, Miloš a Ota NOVOTNÝ. *Business intelligence v podnikové praxi*. Praha: Professional Publishing, 2012. ISBN 978-80-7431-065-2.

RUD, Olivia Parr. 2001. *Data Mining*. Praha: Computer Press, 2001. ISBN 80-7226-577-6.

TVRDÍKOVÁ, Milena. 2008. *Aplikace moderních informačních technologií v řízení firmy: nástroje ke zvyšování kvality informačních systémů*. Praha: Grada, 2008. ISBN 978-80-247-2728-8.

Seznam zkratk


B2B	- Business To Business
BI	- Business Intelligence
CRM	- Customer Relationship Management
DAX	- Data Analysis Expressions
DOLAP	- Desktop OLAP
DSA	- Data Staging Area
DS	- datový sklad
DWH, DW	- Data Warehouse
EAI	- Enterprise Application Integration
EIS	- Executive Information Systems
ERP	- Enterprise Resource Planning
ETL	- Extract, Transform, Load
HOLAP	- Hybrid OLAP
IS	- Informační systém
MOLAP	- Multidimensional OLAP
ODS	- Operativní úložiště dat
OLAP	- On-Line Analytical Processing
ROLAP	- Relational OLAP
SCM	- Supply Chain Management
SQL	- Structured Query Language
SSRS	- SQL Server Reporting Services
SW	- Software
XML	- Extensible Markup Language

Prohlášení o využití výsledků diplomové práce

Prohlašuji, že

- jsem byl seznámen s tím, že na mou diplomovou práci se plně vztahuje zákon č. 121/2000 Sb. – autorský zákon, zejména § 35 – užití díla v rámci občanských a náboženských obřadů, v rámci školních představení a užití díla školního a § 60 – školní dílo;
- beru na vědomí, že Vysoká škola báňská – Technická univerzita Ostrava (dále jen VŠB-TUO) má právo nevýdělečně, ke své vnitřní potřebě, diplomovou práci užít (§ 35 odst. 3);
- souhlasím s tím, že diplomová práce bude v elektronické podobě archivována v Ústřední knihovně VŠB-TUO a jeden výtisk bude uložen u vedoucího diplomové práce. Souhlasím s tím, že bibliografické údaje o diplomové práci budou zveřejněny v informačním systému VŠB-TUO;
- bylo sjednáno, že s VŠB-TUO, v případě zájmu z její strany, uzavřu licenční smlouvu s oprávněním užít dílo v rozsahu § 12 odst. 4 autorského zákona;
- bylo sjednáno, že užít své dílo, diplomovou práci, nebo poskytnout licenci k jejímu využití mohu jen se souhlasem VŠB-TUO, která je oprávněna v takovém případě ode mne požadovat přiměřený příspěvek na úhradu nákladů, které byly VŠB-TUO na vytvoření díla vynaloženy (až do jejich skutečné výše).

V Ostravě dne 25.04.2015


.....
Bc. Adam Rybář

Seznam příloh

Příloha 1: Script pro vytvoření struktury v databázi na produkčním serveru firmy

Příloha 2: Pozadí doplňku Power Pivot

Příloha 3: Ukázka uživatelského prostředí dashboardu

Příloha 1: Script pro vytvoření struktury v databázi na produkčním serveru firmy

```
USE [Evolio_EX_Diagnostics]
GO

--- Vytvoření tabulky pro import kalendáře ---

CREATE TABLE [dbo].[Kalendar]
(
    [Datum] [date] NOT NULL PRIMARY KEY,
    [Den] [int] NOT NULL,
    [Týden] [int] NULL,
    [Měsíc] [int] NULL,
    [Kvartál] [int] NULL,
    [Rok] [int] NULL,
    [Den_v_roce] [int] NULL,
    [RokTýden] [int] NULL,
    [RokMěsíc] [int] NULL,
    [RokKvartál] [int] NULL,
    [Den_v_týdnu] [int] NULL,
    [JménoDne] [varchar](50) NULL,
    [JménoMěsíce] [varchar](50) NULL,
    [Den_v_měsíci] [varchar](50) NULL,
    [FormatRRRRMMDD] [varchar](50) NULL,
    [FormatRRRR_MM_DD] [varchar](50) NULL,
    [KvartálRok] [varchar](50) NULL
)

--- Nahrát kalendář ze souboru ---

--- Přidání sloupce Datum, aby bylo možné propojení s kalendářem ---

GO
ALTER TABLE [dbo].[Profylaxe_201_AddinReport1]
ADD Datum date
GO
UPDATE [Evolio_EX_Diagnostics].[dbo].[Profylaxe_201_AddinReport1]
SET Datum = CAST(
    CAST(year AS VARCHAR(4)) +
    RIGHT('0' + CAST(month AS VARCHAR(2)), 2) +
    RIGHT('0' + CAST(day AS VARCHAR(2)), 2)
    AS DATE)
GO
```

```

-----
GO
ALTER TABLE [dbo].[Profylaxe_210_EvolioLoginsReport]
ADD Datum date
GO
UPDATE [dbo].[Profylaxe_210_EvolioLoginsReport]
SET Datum = CAST(
    CAST(year AS VARCHAR(4)) +
    RIGHT('0' + CAST(month AS VARCHAR(2)), 2) + '01'
    AS DATE)
-----
GO
ALTER TABLE [dbo].[Profylaxe_220_DataboxReport]
ADD Datum date
GO
UPDATE [dbo].[Profylaxe_220_DataboxReport]
SET Datum = convert(date,StartTime)
-----
GO
ALTER TABLE [dbo].[Profylaxe_530_LustrationReport]
ALTER COLUMN Datum date
GO
-----
--- Přidání cizího klíče kvůli propojení s kalendářem do faktových tabulek ---
-----
ALTER TABLE [dbo].[Profylaxe_201_AddinReport1] WITH CHECK ADD CONSTRAINT [FK_Profylaxe_201_AddinReport1_Kalendář]
FOREIGN KEY ([Datum]) REFERENCES [dbo].[Kalendář]([Datum])
ALTER TABLE [dbo].[Profylaxe_201_AddinReport1] CHECK CONSTRAINT [FK_Profylaxe_201_AddinReport1_Kalendář]
--ALTER TABLE [dbo].[Profylaxe_201_AddinReport1] DROP CONSTRAINT [FK_Profylaxe_201_AddinReport1_Kalendář]
-----
ALTER TABLE [dbo].[Profylaxe_210_EvolioLoginsReport] WITH CHECK ADD CONSTRAINT [FK_Profylaxe_210_EvolioLoginsReport_Kalendář]
FOREIGN KEY ([Datum]) REFERENCES [dbo].[Kalendář]([Datum])
ALTER TABLE [dbo].[Profylaxe_210_EvolioLoginsReport] CHECK CONSTRAINT [FK_Profylaxe_210_EvolioLoginsReport_Kalendář]
--ALTER TABLE [dbo].[Profylaxe_210_EvolioLoginsReport] DROP CONSTRAINT [FK_Profylaxe_210_EvolioLoginsReport_Kalendář]
-----
ALTER TABLE [dbo].[Profylaxe_220_DataboxReport] WITH CHECK ADD CONSTRAINT [FK_Profylaxe_220_DataboxReport_Kalendář]
FOREIGN KEY ([Datum]) REFERENCES [dbo].[Kalendář]([Datum])
ALTER TABLE [dbo].[Profylaxe_220_DataboxReport] CHECK CONSTRAINT [FK_Profylaxe_220_DataboxReport_Kalendář]
--ALTER TABLE [dbo].[Profylaxe_220_DataboxReport] DROP CONSTRAINT [FK_Profylaxe_220_DataboxReport_Kalendář]
-----
ALTER TABLE [dbo].[Profylaxe_530_LustrationReport] WITH CHECK ADD CONSTRAINT [FK_Profylaxe_530_LustrationReport_Kalendář]
FOREIGN KEY ([Datum]) REFERENCES [dbo].[Kalendář]([Datum])
ALTER TABLE [dbo].[Profylaxe_530_LustrationReport] CHECK CONSTRAINT [FK_Profylaxe_530_LustrationReport_Kalendář]
--ALTER TABLE [dbo].[Profylaxe_530_LustrationReport] DROP CONSTRAINT [FK_Profylaxe_530_LustrationReport_Kalendář]
GO
-----

```


Příloha 2: Pozadí doplňku Power Pivot

PowerPivot for Excel - DASHBOARD_PowerPivot_0.5.xlsx									
<div><div><div>Domů</div><div>Návrh</div><div>Rozšířené</div></div><div><div><div><div><div><div></div><div></div></div><div><div>Odstranit</div><div>Ukotvit</div></div><div><div>Šířka</div><div>Sloupce</div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><div></div></div><div><div></div><div></div></div></div><div><div><div><div><div></div><</div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div>									

Příloha 3: Ukázka uživatelského prostředí dashboardu

